# 字母位置编码的模型对比及其效应解释

李璜夏[1]，陈新炜[2]，药盼盼[1]

1 北京语言大学心理学院，北京，100083

2 北京外国语大学中国语言文学学院，北京，100089

**摘 要** 在视觉词汇识别过程中，字母的位置信息发挥了重要的作用。过去几十年间，关于字母位置编码的研究极大地推动了各种理论框架的发展，这些理论旨在解释不同的实验效应及其背后的认知加工机制。文章系统介绍了关于字母位置编码的六个理论模型，包括重叠模型(the Overlap Model)、开放双字母组模型(the Open-Bigram Model)、序列编码模型(the SERIOL Model)、空间编码模型(the Spatial Coding Model)、贝叶斯读者模型(the Bayesian Reader)以及N-字母组位置编码模型(PONG: the Positional Ordering of N-Grams)。这些模型涵盖了从重叠编码到序列和空间编码等不同的认知加工机制，代表了字母位置编码领域中的重要理论框架。文章从模型结构、理论基础、词汇识别逻辑、跨语言适应性解释以及常见效应解释等方面进行对比分析，并且对模型尚未能解释的效应进行了总结。基于对这些模型的分析总结，未来模型建构可以整合更多实证研究结果以及不同类型的实验数据，以增强模型解释力度。此外，考虑到跨语言因素以及第二语言的研究成果，探究字母位置加工及相关模型的跨语言一致性将是一个有价值的研究方向。

**关键词** 字母位置编码模型，词汇识别，转置效应，模型对比

收稿日期：2024-12-17；通讯作者：药盼盼（yaopp@blcu.edu.cn）；基金：中外语言交流合作中心国际中文教育研究课题（23YH33B）。

# 1 引言

视觉词汇识别主要涉及对字母身份信息(identity information)和位置信息(position information)的加工。身份信息指视觉对象的身份，即字母本身所代表的意义，例如“CRASH”一词中“C”、“R”、“A”、“S”、“H”这五个字母本身就是其身份信息。位置信息(即字母位置)指视觉对象在字母串中所在的位置，如“CRASH”一词中“R”所在的位置是 2，或是“C”和“A”的中间(Grainger, 2008, 2018)。字母身份信息对词汇识别的影响显而易见，前人研究表明字母位置信息同样对视觉词汇识别有重要影响(Davis, 2010; Johnson et al., 2007; Logan, 2021; Whitney et al., 2012)。例如，“STOP”、“POST”、“SPOT”含有完全相同的字母，但是各个字母位置不同，导致词汇含义完全不同。这些现象引发了研究者们对人脑如何编码字母位置信息、字母位置如何影响词汇识别以及句子阅读等问题的探究，并由此构建了一系列的模型对此现象背后的认知机制进行模拟和解释。

基于不同的理论流派，研究者们建构了不同的模型框架，主要包含严格加工和灵活加工这两种。早期模型基于槽的编码(slot-based encoding)假设字母位置加工较为严格。具体而言，字母位置在词汇加工的早期被编码，词汇中的每个字母独立于该词汇中的其他字母被标记在字母串的特定位置,字母位置的变化会显著影响词汇识别。此类模型包括交互激活模型(Interactive Activation Model, McClelland & Rumelhart, 1981)、多重读出模型(the Multiple Read-out Model, Grainger & Jacobs, 1996)、双路径联级模型(the Dual-Route Cascaded Model, Coltheart et al., 2001)、交互验证模型(the Activation-verification Model, Paap et al., 1982)。然而一些研究表明，在阅读过程中字母位置编码并非如此严格，大脑以一定程度的灵活性或不确定性对字母位置进行编码(Kirkby et al., 2022; Lupker et al., 2008; Marcet et al., 2019; Pagan et al., 2016; Perea et al., 2012, 2018; Perea & Acha, 2009; Perea & Carreiras, 2006; Perea & Lupker, 2004; Snell et al., 2018, 2022; Snell, 2024)。基于这些发现，一些模型致力于解释字母位置编码灵活性背后的认知机制。其中重叠模型(the Overlap Model, Gomez et al., 2008)和贝叶斯读者模型(the Bayesian Reader, Norris, 2006; Norris et al., 2010; Norris & Kinoshita, 2012)基于噪声通道编码理论建构；空间编码模型(the Spatial Coding Model, Davis, 2010)基于空间编码理论建构；开放双字母组模型(the Open-Bigram Model, Grainger & Heuven, 2003)和序列编码模型(the SERIOL Model, Whitney, 2001)基于相对位置编码理论建构；N-字母组位置编码模型(PONG: the Positional Ordering of N-Grams, Snell, 2024)基于分割加工理论建构。

迄今为止，针对字母位置编码已有大量研究，并且涵盖了不同的语言类型(包括一些字符语言，如汉语)。目前已有文献对字母位置编码的研究进行了总结分析，但缺乏对相关模型的对比介绍。对字母位置编码的相关模型进行系统对比，可以加深对字母位置编码认知机制的了解。本文将系统介绍与字母位置灵活编码相关的六个经典模型，包括重叠模型(the Overlap Model)、开放双字母组模型(the Open-Bigram Model)、序列编码模型(the SERIOL Model, 以下简称 SERIOL 模型)、空间编码模型(the Spatial Coding Model)、贝叶斯读者模型(the Bayesian Reader)以及 N-字母组位置编码模型(PONG: Positional Ordering of N-Grams, 以下简称 PONG 模型)，对比模型的结构、逻辑、理论基础以及各个模型对常见实验效应的解释，并基于此对视觉词汇识别中字母位置加工模型的未来发展趋势进行展望。为了系统直观的对比这六个经典模型，各个模型的特点总结在表 1 中。

# 2 模型结构概览

## 2.1 开放双字母组模型(The Open Bigrams Model)

开放双字母组模型由 Grainger 和 van Heuven 等人基于英文阅读的研究结果提出(Grainger & van Heuven, 2003; Grainger & Whitney, 2004)。该模型详细定义了字母加工以及词汇表征的过程，模型结构如图 1 所示。在开放双字母组模型中，视觉词汇的识别包含三个层级：字母数组(letter array)、相对位置地图(relative position map)、整词正字法表征(O-words)。第一层为字母数组，由专门为阅读开发的字母特征(包括空格)检测器组成。这些特征检测器可以同时加工有限长度的字母串中的所有字母。每个字母检测器均能提供有关其在特定视网膜位置的字母身份信息。以字母串“SILENCE”为例，字母数组根据视觉特征的输入可以检测到“#”、“S”、“I”、“L”、“E”、“N”、“C”、“E”、“#”(“#”表示词边界)等身份信息。这些信息随后被传递到第二层——相对位置地图。在此层级中，结合了来自不同视网膜位置的信息，使用开放双字母组表征字母串中各个字母的相对位置。为简化双字母组的激活随着组成字母之间的距离增长而持续减少的情况，模型假设双字母组的最大间隔为两个字母。例如，字母串“SILENCE”在这一层级可以激活 15 个双字母组，包括“SI”、“SL”、“SE”、“IL”、“IE”、“IN”、“LE”、“LN”、“LC”、“EN”、“EC”、“EE”、“NC”、“NE”和“CE”。该模型假设该层级中所有激活的双字母组激活值均为 1，而其他未包含在内的双字母组激活值为 0。这些激活的开放双字母组被传递到第三层级——整词正字法表征。这一层级包含所

有已知的词汇，通过与相对位置地图中的开放双字母组是否兼容进行双向促进和抑制的交互作用，以及通过词汇之间的横向抑制，实现最终的词汇识别。例如，双字母组“CE”的输入可以促进“SILENCE”、“SCIENCE”、“SLICE”等包含相应字母组词汇的激活，同时这些词汇也会向双字母组“CE”输入反馈激活。词汇层的“SILENCE”、“SCIENCE”和“SLICE”等词相互竞争，最终激活值最高的词汇“SILENCE”被识别。

在开放双字母组模型的理论框架中，相对位置编码通过将特定字母标记为字母串中的起始位置(通常是通过检测字母串的第一个有效字母信号以及特殊的符号来辅助标记)，将另一个字母标记为字母串的结束位置(通过检测字母串末尾的字母或特定的终止符号实现)，并通过将最少量的相对位置信息分配给内部字母来进行加工。模型重点关注字母间的相对位置关系而非每个字母的精确位置，从而大幅减少加工的复杂度并提高了加工效率。

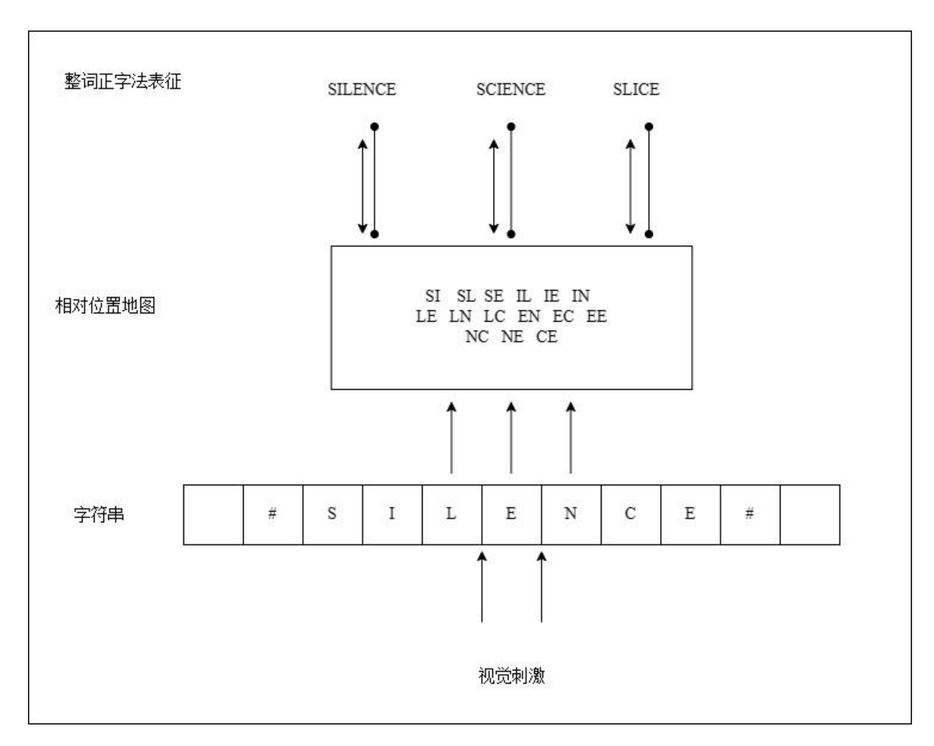



图 1 *开放双字母组结构示意图(原图参见 Grainger & Heuven, 2003)*

## 2.2 SERIOL 模型(Sequential Encoding Regulated by Inputs to Oscillations within Letter Units)

SERIOL 模型由 Witney 等人提出，旨在解释字母位置是如何编码的(Whitney, 2001; Whitney, 2004a; Whitney & Berndt, 1999)。SERIOL 模型沿用了开放双字母组表征字母位置信

息，但在此基础上还加入了连续时间激活模式与开放双字母组相关联(Davis & Bowers, 2006)。SERIOL 包括五个层级，自下而上分别为视网膜层级、特征层级、字母层级、开放双字母组层级和词汇层级，模型结构如图 2 所示。

SERIOL 模型假设每个层级具有不同的激活动态，并使用“节点”来指代基本计算单元。在视网膜层级，一个节点代表一个像素。该层级中的节点根据来自外部的视觉输入在视网膜上进行精确的拓扑组织。由于视锥细胞在视网膜上的分布特征(视觉中央凹分布最多)，感知敏锐度随着与注视点距离的增加而减弱，因此在该层级中形成以注视点为中心向左右两边逐渐减弱的激活梯度。视网膜层级到特征层级的连接对应于从视网膜到大脑半球的信息流，由此将激活梯度输入特征层级。在特征层级的节点识别亚正字法特征，通过模拟大脑左右半球在提取特征时的独立性和语言加工的优势半球导致的差异性，结合阅读方向将感知敏锐度梯度转化为位置信息梯度。具体而言，该模型假设在左视野(右半球)中，敏锐度梯度的斜率在激活字母特征时被反转，而在右视野(左半球)中则保持不变。斜率反转机制是通过两个途径实现的：在左视野(右半球)中自下而上的促进连接和特征层内从左到右的抑制连接上设置更大的权重。该模型假设读者在阅读习得过程中可以学到这些大脑半球间不同的权重和抑制的方向性。在胼胝体转移之后，组合两个部分的梯度，形成一个从左到右单调递减的位置梯度。在字母层级，节点识别各个字母。特征层级的激活梯度引起跨字母节点的时间激活模式，字母的相对位置通过与其他字母节点的激活时间差来确定，即来自特征层级的输入被转换为跨节点的时间序列激活模式，通过激活的先后顺序标记位置信息。这种时间激活模式通过与神经元内的亚阈值振荡(即神经元电位的周期性波动)交互实现。根据低频(5-12Hz)内的 40Hz 子周期是视觉语言加工的基础，模型假设每个字母位置对应于大约 200ms 振荡周期内的一个连续 25ms 子周期。例如，如果一个字母位置在 200ms 的周期内的前 25ms 内被激活，这表明它是字母串中最先被识别的字母；如果是在接下来的 25ms 内激活，那么它是第二个被识别的字母，以此类推。通过这种方式，大脑可以利用神经元的时间响应特性来精确地编码和区分字母序列中每个字母的位置。此过程中的横向抑制可以确保一次仅触发一个字母节点。在开放双字母组层级，模型提供了一种将时间表征解码为非时间编码的机制——有序的字母对。具体来说，双字母组节点的激活取决于包含其组成字母的字母节点的激活水平(随着输入水平的增加而增加)和字母节点激活之间的时间间隔(时间间隔增加，激活水平降低)。例如，在字母串“CART”中，字母层级的输入可以激活“#”、“CA”、“AR”、“RT”、“CR”、“AT”、“CT”和“#”，其中“CA”的激活程度高于“CR”。在词汇层级，节点的输入是开放双字

母组的激活加权之和，此节点将识别整个词汇包含的双字母组集合。“CART”通过这些激活的双字母组成为最活跃的词汇节点，从而完成对词汇的整体识别。

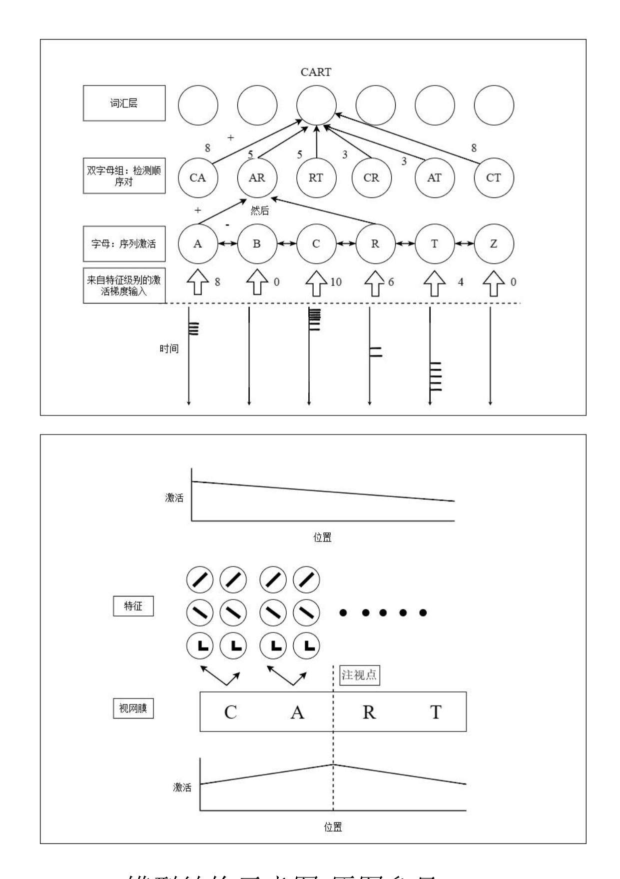



图 2 *SERIOL 模型结构示意图(原图参见 Whitney, 2001)*

## 2.3 空间编码模型（the Spatial-coding Model）

Davis 等人提出空间编码模型来解释字母位置的编码过程(Davis & Bowers, 1999，2010)。模型假设每个节点只对应一个唯一的表征：包括视觉特征、单个字母或词汇。并且模型中包含了专门用于编码外部字母(首字母和尾字母)和词长的表征层级。该模型采用 Grossberg(1978)提出的空间编码机制解决字母位置编码，协调从字母层到词汇层的信号传输。模型结构如图 3 所示。

空间编码模型假设特征和字母层内的节点代表不同位置特定通道的独立子集。书面文本的输入会激活字母一系列特征。特征层级的节点识别字母各项特征，激活的特征向包含该特征的所有字母节点发送促进信号，并向不包含该特征的所有字母节点发送抑制信号。字母层级的节点表征抽象字母身份。抽象字母身份是指字母的识别和加工不依赖于其在具体视觉场景中的形状、位置或上下文，而是以一种去除具体视觉属性的方式进行识别和表征，即被抽

象化为一个独立的、去除具体视觉和空间属性的“身份”。例如字母“A”无论是打印字体还是手写的潦草字体，都能表示字母“A”。在字母层和词汇层之间存在空间编码器，其主要功能是使用空间相位编码传达位置信息。空间相位编码通过将字母层级输入的信号转换为一个二维坐标并形成正态分布，利用正态分布描述字母的位置。其中分布的中心代表实际位置，宽度则表示位置的不确定性。字母节点以固定的周期输出信号，不同字母节点在周期内的激活时间不同，形成不同的相位。这些相位编码字母的相对位置，例如较早输出的字母节点表示较早的位置。空间编码器从第一个字母通道开始，通过周期信号激活其中的字母节点，标记为相位 1。随后，空间编码器将其“注意力”向右移动到下一个字母通道，通过下一个周期脉冲激活相位为 2 的字母节点。这一过程持续进行，直至空间编码器处理完最后一个字母通道。随后位于字母层和词汇层之间的接收器节点拦截空间编码信号。每个接收器节点通过对输入信号的相位进行移位来计算信号与预期位置之间的差异，这些差异反映输入信号和存储模板之间的匹配程度。最后采用“winner-take-all”的原则将激活值最大的节点信号输入到词汇层。接收器输出的信号在词汇节点处进行整合。此外，该模型引入了外部字母库和字段长度模块。外部字母库用于编码首字母和尾字母。这两个库中包含字母表中每个字母的字母节点，且都和词汇层之间存在促进连接。如图 3 所示，首字母库中的“C”节点和尾字母库中的“T”节点都向“CAT”词汇节点发送促进连接。空间编码模型认为，基于字母层的总激活和独立的视觉输入信号，在加工早期就可以利用有关刺激长度的信息，因此字段长度模块内节点的功能是编码当前输入刺激的长度，即词长。在词汇层，节点通过接收器的输入、外部字母库和词长模块的信号输入以及横向抑制机制，最终决定词汇节点的激活。

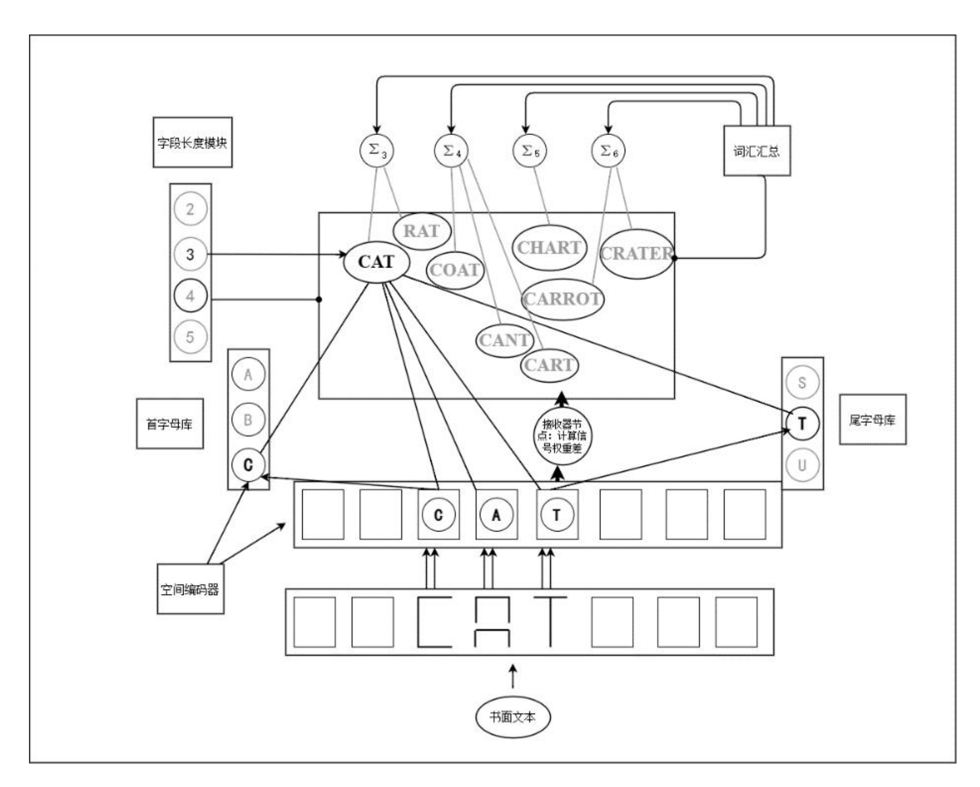



**图** 3 *空间编码模型位置编码示意图(原图见 Davis, 2010)*

## 2.4 重叠模型（The Overlap Model）

重叠模型由 Gomez 等人(2008)基于位置噪声编码理论建构。不同于其他视觉词汇识别模型，重叠模型没有明确的结构框架，模型对字母位置的表征如图 4 所示。在重叠模型中，一个字母串在不加掩蔽且不受时间限制的条件下呈现时，可以产生准确的位置编码；而短暂呈现的字母串在字母位置上有不同的分布(这种位置上的分布仅代表阅读中的初始编码过程)。具体而言，对于一个字母串中的任意一个字母，假设在初始加工阶段其至少与一个以上的位置相关联。例如，对于字母串“TRAIL”，字母“A”将与位置 3 相关联，但是根据标准差的大小(模型通过自由参数设置每个字母位置不同的标准差)，它也会在一定程度上与位置 2 和 4，甚至位置 1 和 5 相关联。在模拟经典实验效应时，重叠模型通过计算两个字母串之间的重叠度来衡量拼写相似度。重叠度的计算机制如下：对于每个位置，将测试字母串对应槽中的字母曲线下的面积乘以学习字母串中相同槽的字母的面积，并将这些乘积在所有槽上求和。模型计算目标刺激与正确备选项和干扰项之间的重叠量，然后使用幂函数将这种重叠度转换为正确反应比例。幂函数中的缩放参数使得较小的重叠差异能够产生较大的准确性。如图 4 的第二个图，“TRAIL”及其转置词汇“TRIAL”，在位置 3 和 4 中的重叠，具体取决于第三和第四字母位置的标准差。而图 4 的第一个图表示“TRAIL”和“TRAIN”第五个位置

没有字母匹配。

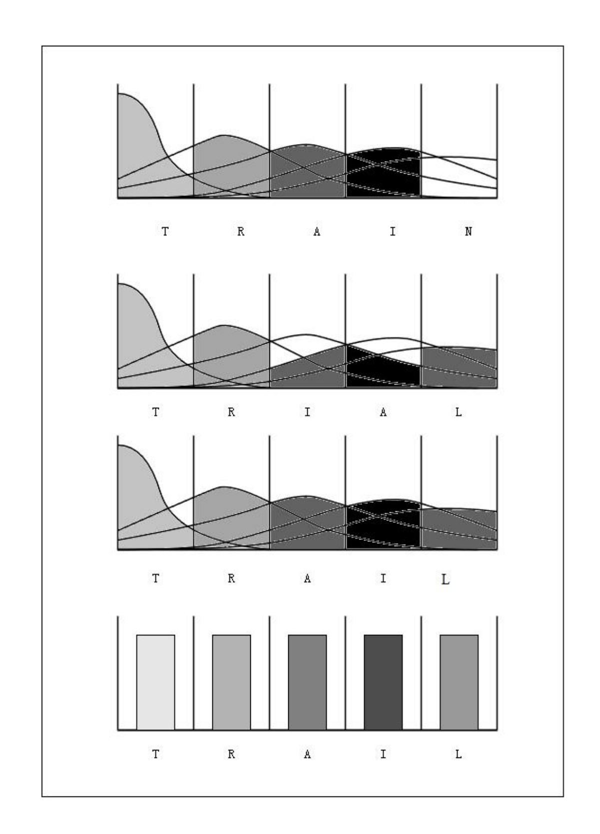



图 4 *根据重叠模型的字母位置编码表征(原图参见 Gomez et al., 2008)*

*注.* 图中不同程度的灰色表示不同的字母。

## 2.5 贝叶斯读者模型（The Bayesian Reader）

另外一个基于噪声通道编码的模型是 Norris 等人(2012)提出的加入噪声的贝叶斯读者模型。该模型基于英语、法语和荷兰语等多语种数据建构，扩展了 Norris 等人在 2006 年提出的第一版贝叶斯阅读者模型(Norris et al., 2006)以及 2010 年的刺激采样理论(Norris et al., 2010)。引入噪声通道的贝叶斯读者模型假设读者可以在词汇识别过程中根据噪声通道的输入做出接近最优的决策。由于存在噪声，字母身份和字母位置存在不确定性，该模型的任务是充分利用噪声通道中的可用信息来推断最有可能被识别的词汇。模型结构如图 5(a)所示。

建立于噪声通道的贝叶斯读者模型没有明确的模型结构，而是更偏向于计算词汇出现的可能性。该模型假设每个字母向量有 27 个坐标（26 个字母和一个空字母），每个字母通过将其中一个元素设置为 1，其余设置为 0 来表示。模型输入使用固定长度的窗口，设置为 12 个字母(所有输入字母串从窗口的第二个位置开始以满足之后字母身份信息从相邻位置泄漏

的假设，其他位置填充空字母)。为了更好地模拟人类视觉的不确定性以及在面对字母位置错误或位置模糊时的阅读能力，模型假设视觉输入经过一个带有噪声的通道，这个通道可能会引入字母身份和位置的混淆，甚至可能导致字母的插入或删除。位置噪声的概念允许字母身份信息在相邻位置之间"泄漏"。在该模型中，字母的位置和身份的识别是一个动态的过程，依赖于累积证据和不断更新的概率分布。在初始状态下，每个字母在特定位置上的身份和位置并不确定，因此模型为每个位置和字母分配一个初始概率分布。随着时间的推移和采样次数的增加，模型从视觉输入中获取更多的特征信息，并利用这些观测到的特征来更新每个位置上字母的概率分布。这个过程使用贝叶斯定理来计算。具体而言，每个位置 i 上都有一个字母 c 出现的先验概率 $P(c_i)$。当模型接收到视觉输入时，会从中提取观测特征 E。对于每个可能的字母 c 在位置 i 上，模型需要计算给定该字母在该位置上时，观测到当前输入特征 E 的概率，即可能性 $P(E|c_i)$。这一步的目的是衡量当前输入特征与每个可能字母在每个位置上的匹配程度。然后使用贝叶斯定理，将先验概率 $P(c_i)$与可能性 $P(E|c_i)$结合，计算后验概率 $P(c_i|E)$，即在观测到输入特征 E 后，字母 c 在位置 i 上出现的更新后的概率。随着更多证据的累积，这些后验概率会逐步更新，字母身份和位置的不确定性逐渐减少。当证据积累足够多时，所有概率分布收敛于一个最有可能的字母序列。模型的噪声采样如图 5(b)所示：早期(较平坦的曲线)字母“O”出现在字母“W”或“R”的位置上的可能性仍然很大，而到了晚期(高窄曲线)这种不确定性降低。在确定可能的字母序列之后，贝叶斯读者模型使用加权编辑距离方法来处理字母序列与模型词表中词汇之间的差异。编辑距离(edit-distance metrics)(Levenshtein et al., 1966; Damerau, 1964)指将一个字母串转换为另一个字母串所需的基本操作(插入、删除和替换等)的数量。例如，将“WORD”转换为“WORLD”需要插入一个字母，编辑距离为 1，再由“WORLD”转换为“WORLLD”需要再添加一个字母，此时编辑距离为 2，以此类推。模型会通过字母概率以及删除和插入概率的乘积计算可能的编辑路径以估算每个词汇的可能性，最后选择最优路径作为最终的词汇识别结果。

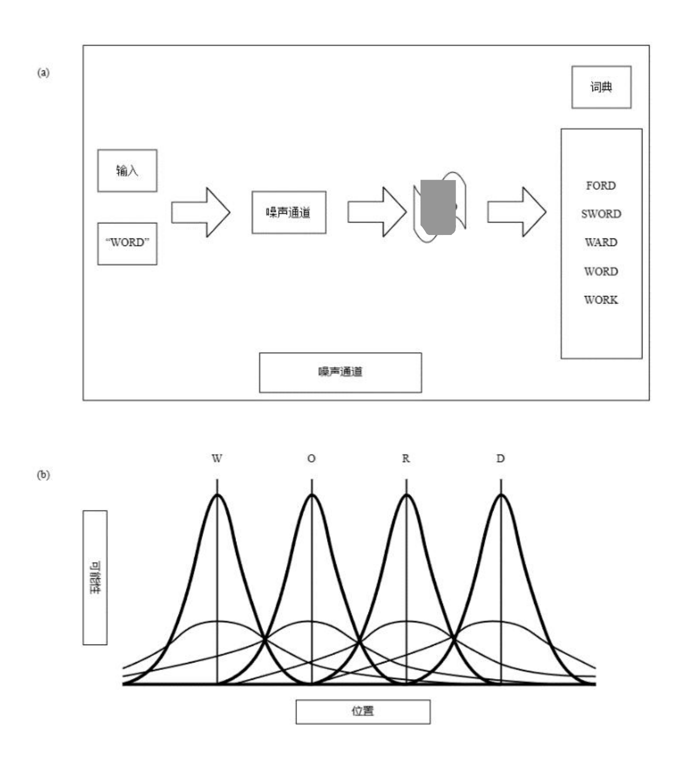



**图 5** *加入噪声通道的贝叶斯读者模型(原图详见 Norris, 2006; Norris & Kinoshita, 2012)*

*注.* (a)为噪声通道；(b)为噪声采样示例，表示时间上早期(平缓曲线)和晚期(高窄曲线)的位置不确定性。

## 2.6 N-字母组位置编码模型(PONG: the Positional Ordering of N-Grams)

最近，Snell 提出了一种关于大脑如何计算字母位置的新模型——N-字母组位置编码模型(PONG: the Positional Ordering of N-Grams, Snell, 2024)。PONG 模型包含两个重要假设：一是大脑通过编码 N-字母组来进行词汇识别；二是字母位置通过大脑两半球激活差异来估计。PONG 模型的核心是引入新的 N-字母组(N-字母组类似于双字母组，其中 N≥1)作为表征中介，来处理字母和字母之间的位置关系。模型结构如图 6 所示。

在完整框架之外，模型额外设置了一个学习者模块(模型原文未在框架图中绘制)。学习者模块在输入开始加工之前运行，考虑了读者的先前经验。具体而言，该模块根据语言输入进行训练，例如文本材料或词汇表，从中得出词汇(fw)和 N-字母组(fn)的统计频率。在这个部分，模型学习到的 N-字母组是训练材料中不被词间空格间隔的所有连续字母组合。例如，从语料“I do”中，PONG 模型可以学习到的 N-字母组包括“I”、“D”、“O”和“DO”，但不包括“ID”或“IDO”。除了计算统计频率外，学习模块还包括为所有已知词汇创建空间编码，将所有组

成 N-字母组的左右位置转换为-1 到 1 的偏侧性范围。例如对于词汇“right”来说，包含的 N-字母组及其偏侧性为：“R”(-1)、“RI(-0.75)”、“RIG(-0.5)”、“I(-0.5)”、“RIGH(-0.25)”、“IG(-0.25)”、“RIGHT(0)”、“IGH(0)”、“G(0)”、“IGHT(0.25)”、“GH(0.25)”、“GHT(0.5)”、“H(0.5)”、“HT(0.75)”、“T(1)”。

PONG 模型的核心结构一共包含了四个层级：(1) 视觉和注意力、(2) N-字母组激活、(3) N-字母组偏侧性估计和(4)词汇激活。在视觉和注意力层，视觉模块为后续 N-字母组的激活提供视敏度信息。视敏度的计算主要由注视点的离心率决定。除此之外，视觉模块可以提供词长信息，在之后的词汇激活中会根据备选词与预期词汇长度的匹配程度进行加权。该层级另外一个部分为注意模块，为 N-字母组提供注意权重信息。与视觉模块一致，注意焦点集中在一个字母位置上。PONG 假设注意根据以注意焦点为中心的正态分布调节输入，注意权重针对每个半球单独计算，由注意力焦点的离心率(与视觉模块计算一致)和注意宽度(设定为 4°视角)决定。在 N-字母组激活层，所有存在的 N-字母组的激活都是视敏度和注意的函数。如果一个 N-字母组在视野中多次出现，则对该 N-字母组检测器的所有激活进行求和。在 N-字母组偏侧性估计层级，通过偏侧性对 N-字母组的空间位置进行精确的判定。偏侧性的计算采用最小激活检测法。首先计算 N-字母组在左半球和右半球的检测器激活值($A_{left}$和$A_{right}$)。当 $A_{left} \neq A_{right}$ 时，偏侧性通过取两个半球中检测器的左右偏性的平均绝对值，并根据左半球激活值是否小于右半球激活值将其设为正或负来进行计算。当 $A_{left} = A_{right}$ 时，偏侧性为 0。这一计算过程具有一定的不确定性，但模型通过正态分布采样加以处理。此外，所有 N-字母组的偏侧性之和可以决定整个词汇的偏侧性，词汇的整体偏侧性反过来又作为一个锚定点，可以根据这一锚定点调整和纠正各个 N-字母组的相对偏侧性，从而确保 N-字母组偏侧性不受词汇位置的影响，可以与词汇表征匹配。最后在词汇激活层级，词汇的激活程度取决于(a) N-字母组是否出现在词汇中，(b) 该词汇与预期词汇长度的匹配程度，(c) N-字母组的估计偏侧性有多大程度匹配学习到的偏侧性、(d) 词汇的频率、(e) N-字母组的频率和(f)N-字母组的激活值。

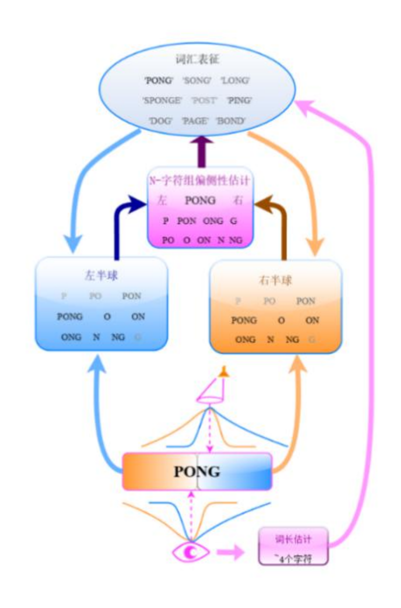



图 6 *PONG 模型结构示意图(原图参见(Snell, 2024))*

*注.* 眼睛注视位置由眼睛符号表示，视觉空间注意力由灯光符号表示。橙色和蓝色曲线表示每个半球的激活作为离心率的函数，橙色代表大脑右半球，蓝色代表大脑左半球。

# 3 模型对比

## 3.1 结构

从总体来讲，上述六个模型都涉及词汇识别模块。但各个模型在结构上有所差异。重叠模型和贝叶斯读者模型基于噪声通道编码字母身份和位置，没有引入专门的正字法表征或字词识别机制，而是用噪声采样表征字母加工机制。其他模型大致都包含了视觉层、字母层、词汇层，但在细节方面也稍有不同。在开放双字母组模型和 SERIOL 模型中，字母层与词汇层之间通过双字母组层表征字母的相对位置。在 PONG 模型中，双字母组层级被 N 字母组和 N 字母组偏侧化估计层所代替，通过 N 字母组在不同大脑半球的激活值来标记字母的位置。在空间编码模型中，字母层与词汇层之间通过使用空间编码器编码字母位置。另外该模型还包含了一个首字母库和一个尾字母库，并通过促进性连接与词汇层相连，用于编码词汇的首尾字母。除此之外，在空间编码模型和 PONG 模型中都包含一个词长估计模块用以估计词长，进而与词典中的词汇进行匹配。PONG 模型还考虑到先验经验的影响，在模型之外设置了学习者模块。六个模型中，SERIOL 模型对低级视觉加工有更为详细的设置：在特征层级之前包含视网膜层级用以激活低级视觉信息。

## 3.2 理论基础

上述六种模型基于四种不同的理论流派：基于噪声槽位编码、空间编码、相对位置编码和分割加工理论。每一种理论根据其假设对字母的位置灵活性编码做出了不同的贡献。

### 3.2.1 噪声槽位编码理论

McClelland 和 Rumelhart(1981)提出，给定位置的视觉输入可能激活周围位置的探测器，这意味着大脑必须考虑字母位置不确定性或存在“噪声”。根据噪声编码理论，字母位置不是一个点，而是在空间中的一个分布。在最强灵活性假设下，即使是位置分布的最远端，其激活概率也不等于零而是趋近于零，表明字母在某种程度上会激活词汇而与位置无关。基于噪声槽位的编码原理在重叠模型以及其贝叶斯读者模型中得到应用。但这两个模型在应用噪声编码理论时也存在一些差异。重叠模型认为位置分布的宽度是一个关键参数，可以决定给定位置处的字母是否激活相邻位置或更远的位置；而每个位置的曲线高度反映了该字母存在的可能性。该假设在多项研究中得以证明(Delooze et al., 2022; Grainger et al., 2014; Snell et al., 2018a)。最新版的贝叶斯读者模型(Norris et al., 2012)沿用了噪声通道编码字母身份和位置的假设，同时添加了编辑距离以解决插入和删除字母的问题。

### 3.2.2 空间编码理论

空间编码最初受到 Grossberg(1978)提出的视觉刺激空间位置的机制的启发，并由 Davis(1999, 2010a)在视觉词汇识别的计算模型中完全实现。空间编码理论提供一种不同于传统槽位编码的视角：假设字母的识别并不依赖于固定位置，而是基于字母在字母串中的相对位置进行匹配。该理论认为视觉输入以“上下文无关”的方式与词汇表征进行匹配。具体而言，视觉输入被看作是一系列快速的扫描，逐个激活每个字母的检测器。通过这种方式，模型学习到每个字母的位置，并根据这些位置来激活词汇。词汇的激活程度取决于输入字母的顺序和位置与已知词汇模式的匹配程度。也就是说，无论字母出现的位置或周围的上下文如何，当输入刺激包含字母“A”时，编码“A”的节点就会被激活。字母串中字母的相对位置由字母节点之间活动的相对模式编码，不同的字母位置会导致不同的空间活动模式(因此称为“空间编码”；“空间”一词并非指视觉空间坐标)。在空间编码模型中，该理论被纳入以解释字母位置编码的灵活性。

### 3.2.3 相对位置编码理论

相对位置编码理论的核心假设是大脑通过双字母组的形式表征字母及其相对位置。例如，双字母组"AB"表明在视觉识别中字母"A"位于"B"的左侧。这种编码方式提供了一种适应字母位置编码灵活性的方法。在双字母组中字母间不相邻也能触发双字母组的激活，例如，"JUDGE"可以激活双字母组"JD"。词汇的激活是一系列双字母组激活的函数。由于双字母组的表征在位置上具有不变性，因此在词汇层面上，不涉及字母的绝对位置。使用此类编码方式的两个主要模型是 SERIOL 模型和开放双字母组模型。但在开放双字母组模型中，相对位置的激活为 0、1 编码；而在 SERIOL 模型中其为连续值编码。另外在 PONG 模型中，Snell 等人认为大脑采用相对位置和绝对位置混合编码的方式对字母身份和位置进行加工，这一过程中视觉信息的激活取决于 N-字母组(N 为≥1 的整数)而非开放双字母组。

### 3.2.4 分割加工理论

分割加工理论，通常也称为中央凹分割理论(Ellis & Brysbaert，2010)，其出发点是对人眼和初级视觉皮层的解剖结构的探究。该理论认为视网膜中央凹约 0.5 度视角的区域内，信息会投射到左右两侧半球共同处理外；其余落在视网膜右半部分的视觉信息最初投射到大脑左半球，而落在视网膜左半部分的信息则投射到大脑右半球。实验证据表明词汇激活很大程度上依赖于"语言主导"半球，特别是左侧枕颞沟中的视觉词形区域(Cohen et al., 2000; Dehaene & Cohen, 2011; Kronbichler et al., 2004; Rauschecker et al., 2012)。基于这些信息，分割加工理论认为正字法信息的输入在信息加工流的某个点上在两个半球之间传输，如果这个转换点在大脑表征单个字母的阶段之后，并且每个字母最开始只在一个半球出现，那么通过视觉输入的半球分割就可以知道哪些字母位于左侧，哪些字母位于右侧。PONG 模型基于分割加工理论建立了模型的核心假设，即模型通过左右半球激活差异估计字母位置，结合 N-字母组合的不同半球表征来解决字母位置的编码问题。

## 3.3 词汇识别

上述六个模型对如何实现词汇识别存在不同的构建。开放双字母组模型认为，词汇的激活依赖于双字母组与词汇节点之间相互促进和抑制的连接。在该模型中，词汇识别不仅依赖于单个字母的位置和顺序，还依赖于这些字母的组合形式。通过双字母组的促进和抑制的加

权综合，最终决定词汇的激活和识别。在重叠模型和贝叶斯读者模型中，都是通过计算字母串之间的相似度来进行词汇识别。其中，重叠模型的词汇识别依赖于整个词的模糊匹配，即重叠度。读者在识别词汇时，会将输入的模糊字母序列与词汇表中的词进行对比，由于字母表征的模糊性，模型允许部分匹配，即使输入字母序列与目标词并不完全一致，词汇识别也可能成功。贝叶斯读者模型则认为词汇识别是基于贝叶斯推理的过程。模型计算每个词汇在给定输入字母序列下的后验概率，其中后验概率最大的词汇被识别。此外，模型会同时结合先验知识(如词频、上下文信息)来更新和调整对词汇的识别概率。在 SERIOL 模型中，词汇由跨上下文单元的激活模式来表示。该模型采用传统的神经网络方法来构建连接，每个词汇节点的输入由双字母组层级的加权激活之和决定。当该词汇跨双字母组节点进行表征时，给定双字母组节点到给定词汇节点的连接权重与双字母组的激活成正比。因此，一个词汇的识别依赖于它所包含的双字母组的激活强度及其组合方式。空间编码模型中词汇识别的核心是叠加匹配算法。空间编码器使用指代相对字母位置的相位代码动态标记字母信号。这些信号由接收器节点拦截以调整相位，解决由重复字母导致的信号竞争，然后接收器输出的信号在词汇节点处整合，实现叠加匹配算法。此外来自外部字母库的输入也有助于词汇节点计算匹配值。除了匹配值之外，词汇节点还计算表示输入刺激与模板之间不匹配的项。通过将这些自下而上的匹配和不匹配的信号、横向抑制和促进的信号，以及来自词长模块的长度匹配或不匹配信号相结合来计算词汇节点的输入，实现最终的词汇识别。在 PONG 模型中，词汇的激活依赖于 N-字母组的激活。这个过程取决于 N-字母组是否出现在词汇中、词汇与预期词汇长度的匹配程度、N-字母组的偏侧性与学习到的偏侧性的匹配程度、词频信息、N-字母组的频率和 N-字母组的激活等综合信息输入，通过对这些信息的加权计算进行词汇识别。

## 3.4 跨语言适应性解释

拼音文字由线性排列的字母构成，且字母音位和字形之间存在较强的映射关系。这种透明度使得在词汇加工过程中，读者可以通过单个字母-音素转换过程有效地进行词汇识别，形成“音-形结合”加工的路径(Frost, 1998)。相比之下，汉语是由非线性结构、形义结合的字符构成。汉字通常同时承载语音和语义的信息(Perfetti & Tan, 1998)。由于汉语字符的表意性和复杂的形状结构，在汉语词汇处理过程中，读者需要将字符整体的视觉特征与语义信息相结合(Perfetti et al., 2005)。因此，汉语字符的词汇识别过程更多地涉及到视觉形状的整体加工和直接的语义激活(Zhou & Marslen-Wilson, 1999)。此外，不同于拼音文字，汉语中不存在

词间空格对词边界进行标记。阅读汉语文本时，读者的首要任务是对连续的字符串进行切分，并对切分出的字符组块（词）进行识别和语义通达。汉语拥有超过 6000 个字符，其中大部分具有意义可以作为语素。这些字符/语素既可以独立成词，又可以互相组合成词。这些显著区别于拼音文字的汉语特异性对字母位置编码模型的跨语言适应性提出了挑战。

虽然现有的一些字母位置编码模型可以被用于解释汉语字符转置效应，如开放双字母组模型与 SERIOL 模型，但却不能解释汉语文本特异性导致的汉语字符转置效应与其他语言存在的差异。研究表明，汉语读者在字符位置编码方面相较于拼音文本读者更加灵活。Davis 等人(2008)在研究英语母语读者的极端转置效应时，发现仅在使用前向启动时才会产生启动效应，反向转置启动对英语读者无效。Yang 等人(2019)考察了汉语中的极端转置现象。在实验中，使用四字词语作为目标词，反向书写的词作为启动词(例如"牢补羊亡"启动"亡羊补牢")。结果发现，即使在极端转置情况下，汉语读者仍能表现出启动效应。为了探讨这一启动效应是否基于正字法，Yang(2020)进行了两项实验：同音启动(佟步锁友(tóng bù suoˇ yoˇ u)--有所不同(yǒu suǒ bù tóng))和语义启动。所有实验结果均未发现语素/语义或音节/语音因素导致反向启动效应，进一步确认了正字法编码是极端启动效应的主要原因。这些研究表明汉语字符位置编码比拼音文字中的字母位置编码更加灵活，而这种语言系统导致的词汇加工机制的差异尚不能被现有模型解释。Lally 等人(2020)使用人工语言训练的结果也表明书写系统的设计显著影响了字母位置编码的精确性，强调在不同语言和书写系统中，字母位置编码并非普遍一致，而是可能受特定书写系统结构的影响。这些研究都为构建发展可以解释跨语言差异性的字母位置编码模型提出了需求。

已有的一些考察汉语词汇识别的模型仅在一定程度上可以对字符位置加工做出解释和推论。Taft 等人(1998)提出的多重交互激活模型(Multi-Level Interactive-Activation Framework)对汉语字符结构特性做出了假设。该模型中，字符识别是一个包含多个层次的过程，不只是停留在字符的整体识别上，而是深入到字符的不同成分(如义符和笔画等)。每个层级都独立运作但又互相联系：不同层级之间的信息会相互传递并促进或者抑制各个节点的激活。具体来说，当识别一个汉字时，字符的部件信息会激活字符的整体信息，而字符的整体信息也会反过来影响部件的激活，这就形成了一个双向的、动态的激活过程。在这个模型中，Taft 等人(1998)强调了字符部件对字符识别过程有显著影响，主要考虑了字符部件的频率和结构特征的作用。例如，在汉字中，左右或上下部件的位置会对整体字符的识别产生不同的作用。这一位置敏感性意味着在识别字符时，部件的空间位置信息会直接影响到该字符的处理和识

别难度。另外，汉字部件的出现频率会影响识别过程中相应节点的激活强度：高频部件更容易被识别和激活，而低频的部件则需要更多的认知资源来加工。

Li 等人(2009)针对汉语不存在词间空格的特点提出了一个词汇识别模型。该模型认为汉语中的词切分和词识别是统一的过程。汉语词汇别涉及三个层级(特征层级、字符层级和词汇层级)之间的节点交互作用。具体而言，一个注视点可以接收的视觉输入激活特征层级的对应节点，进而激活字符层级的对应节点。在某一槽位的字符节点之间相互竞争。字符层的激活节点进而又激活词汇层级的对应节点。最终的词汇识别由词汇层级中的各个节点之间的竞争实现：激活程度最高的词汇节点最先被识别并同时完成词切分。此外，由于汉语阅读方向为从左到右，该模型假设左侧的词汇在竞争中具有优势，因此左侧的词汇比右侧的词汇更快加工。由于该模型假定严格的字符位置编码并和槽位对应，字符位置混乱将导致无法激活在词汇层级的相关节点。因此，在该模型的理论框架下，任何情况下转置假词都不会对基词的识别起促进作用。

Gu 等人(2015)考察汉语中转置字符是否跨词边界对词汇识别的影响。使用边界范式，Gu 等人发现在单词条件下(组成四字词的字符无法切分为两个双字词，如“剑拔弩张”)，即使中间两个字符被转置，预视仍能激活基词，转置条件下的目标词阅读时间显著短于替换条件。然而，由于字符位置不精确，基词激活度较相同预视条件下降，导致转置条件下的阅读时间比相同条件略长。在双词条件下(组成四字词的字符可切分为两个双字词，如“庄严肃穆”)，转置假词预视对目标词的加工没有明显的促进作用，导致阅读时间与替换假词预视条件相近。这一结果说明汉语中的字符转置效应受到转置字符是否跨词边界的影响。相较于跨词边界的字符转置，词内部的字符转置更容易引起对目标词的启动。Gu 等人认为这一实验发现不能被 Li 等人(2009)的分词模型解释。但如果在 Li 等人(2009)的分词模型基础上引入与重叠模型类似的假设（假设每个字符与多个位置相关联，并且每个字符与不同位置相关联的程度作为以字符的实际位置为中心的正态分布的函数出现）则有希望解释汉语中灵活的字符位置编码效应。这也为未来模型的发展和优化提出了新的思考方向。

目前还未有模型可以完备解释汉语文本特异性及汉语中的字符位置效应。本文综述的六个模型均是以拼音文本为基础构建的。其中 PONG 模型本质上是一种学习模型，它通过训练获取正字法的统计特性，进而在加工不同语言时可能会有不同的表现。虽然目前 PONG 模型还未在汉语中应用，但其可以成功模拟出英语和希伯来语之间的加工差异。PONG 模型在对希伯来语进行模拟时，该模型表现出的位置灵活性程度小于英语，并且还产生了阅读速

度的总体差异，即希伯来语的词汇识别速度比英语慢，这与以往研究结果一致(Trauzettel-Klosinski & Dietz，2012)。这些结果证实了 Lerner 等人(2014)的主张：即模型的学习能力是解释跨语言差异的关键因素，而不仅仅是依赖于固定的参数设置。PONG 模型在解释字母位置编码中跨语言差异的能力上进一步巩固了假设 N-字母组的优势。相对于 PONG 模型，其他几个模型由于构建较早，未设置学习训练的机制，在跨语言的差异性的解释上相对较弱。其中贝叶斯读者模型虽然基于英语、法语、荷兰语三种语言构建，但是没有纳入解释不同效应的模块或机制，只对这几种语言中的共有效应做了模拟和解释。

# 4 对常见转置效应的解释

## 4.1 经典转置效应

字母位置编码涉及到字母身份和位置加工，转置字母效应的结果表明这两个加工过程是分离的，字母位置信息加工相较于字母身份信息加工更为灵活(Schoonbaert & Grainger, 2004a; Snell et al., 2018; Winskel et al., 2013)。具体来说，前人研究发现转置字母非词(将词汇中的两个字母调换位置，如 jugde-JUDGE)相对于替换字母非词(将词汇中的两个字母替换，如 juyhe-JUDGE)更能促进其基词 JUDGE 的激活，同时两者的促进作用都小于基词本身(Cao et al., 2022; Grainger & Holcomb, 2009a; Gu et al., 2015; Johnson et al., 2007; Perea et al., 2018; Warrington et al., 2019; Zhang et al., 2021)。在开放双字母组模型中，开放双字母组的检测和激活可以解释转置字母效应。比如，对于词汇“WORD”来说，激活的双字母组与其转置词“WROD”激活的 6 个双字母组中的 5 个是相同的，而与其替换词“WAFD”仅有 1 个双字母组重叠。因此，根据开放双字母组模型的假设，转置假词相对于替换假词与基词在编码上有更大的相似性，表现为更强的启动效应。重叠模型中假设字母的表征分布在字母串中的顺序位置上，两个字母串之间的相似度是相同身份的字母之间的重叠程度，并且使用幂函数将重叠度转换为准确度。例如，如果字母串是“JUDGE”，那么字母“D”与位置 3 高关联，但与位置 2 和位置 4 相关联的程度较低，与位置 1 和位置 5 相关联的程度最低。根据重叠模型的假设，“JUDGE”的转置词“JUGDE”相对于替换词“JUKBE”与原词在位置分布上有更多的重叠，更为相似。相对于替换假词，转置假词与原词重叠程度更大，因此识别速度更快。贝叶斯读者模型在解释转置效应时认为模型通过累积输入中的噪声样本进行操作。例如在识别“JUDGE”的字转置假词“JUGDE”时，模型首先计算每个字母在所有可能位置上的概

率。由于字母“G”和“D”的位置存在不确定性，模型会将“JUGDE”中的“G”视为可能出现在“JUDGE”的正确位置上，反之亦然。之后模型通过加权编辑距离计算来估计每个可能词汇的可能性。当计算“JUGDE”与“JUDGE”之间的匹配时，模型会将转置字母视为较小的编辑距离，给予较高的匹配分数。SERIOL 模型中，对转置效应的解释也采用了开放双字母组。来自字母层级中的激活梯度信息会激活不同的双字母组，转置条件相对于替换条件可激活与原词更多匹配的双字母组，从而出现转置效应。空间编码模型中，字母的位置通过其相位进行编码，空间编码确保字母的相对顺序得到有效表示。替换非词通过改变字母的具体内容，扰乱了相位编码的准确性和字母与词汇模板的匹配。而字转置假词仅改变字母的位置，相位编码的相对关系仍然保持相对稳定。空间编码模型通过叠加匹配算法进一步解释字母转置效应。在叠加匹配中，输入代码和学习代码之间的匹配是基于它们共有字母的相对位置进行的。字母转置的情况下，共有字母依然存在，叠加匹配仍然能够产生较高的匹配值。而替换字母则改变了字母身份，导致共有字母减少，匹配值降低。PONG 模型对转置字母效应的解释是，对于词汇“WORD”,替换条件“WATD”仅在对应的半球激活两个适当的 N-字母组“W”和“D”。转置条件“WROD”激活四个匹配的 N-字母组，“W”、“D”、“R”和“O”，其中前两个具有适当的偏侧性，而后两个则稍微偏离从词典中学习到偏侧性。即“W”和“D”的偏侧性为原本的-1、1，而“R”和“O”由原来的 0.5、-0.5 变为-0.5、0.5。完整的刺激“WORD”会激活组成词汇“WORD”的所有 N-字母组，并且都具有合适的偏侧性。这样的字母激活值以及偏侧性值导致在最后计算词汇激活值时相同条件最大，转置条件次之，而替换条件最小。

## 4.2 非相邻字母转置效应

在视觉词汇识别中，转置效应的大小受到字母位置是否相邻的影响。以往的研究表明，即使是非相邻字母的转置，也能触发启动效应，但效应相较于相邻字母转置而言较弱 (Aschenbrenner et al., 2017; C. H. Lee & Taft, 2009; Grainger et al., 2006; Gu et al., 2022; Johnson & Eisler, 2012)。开放双字母组模型、重叠模型、空间编码模型、SERIOL 模型以及 PONG 对此进行了解释。开放双字母组模型认为，无论字母是否相邻，其身份编码不变，关键在于相对位置编码的改变。相对于相邻转置，非相邻转置相对位置跨度更大，导致激活的双字母组与原词的匹配度降低，因此，相邻字母转置的加工速度通常快于非相邻字母转置。重叠模型使用重叠度来解释非相邻转置效应。在所有的字母位置编码中，相邻转置的正字法重叠比非相邻转置更多，因而相邻的转置启动应该比不相邻的转置产生更强的效应。空间编

码模型中字母的位置通过相位编码进行表征，非相邻转置会使字母在空间编码中的位置标记变得不一致，导致字母和词汇模板之间的匹配变得困难，从而影响识别的效率。在 SERIOL 模型中，非相邻转置条件下激活的双字母组多于替换条件，因此也能从产生转置效应。但是相较于转置条件，字母对之间的间隔增加，激活的与原词匹配的双字母组更少，因此产生更小的转置效应。PONG 模型在解释非相邻字母转置效应时认为，字母位置通过 N-字母组的偏侧性来估计。相对于相邻转置，非相邻转置可能导致与原词更大的 N-字母组偏侧性估计偏差。但是无论是相邻转置还是非相邻转置，都比替换条件激活更多的 N-字母组，因此，非相邻转置也可以产生转置效应，但要弱于相邻转置。

## 4.3 字母重复、插入和删除

以往的研究发现字母重复“ROCCK”、删除“ROK”或者插入新的字母“ROCNK”虽然改变了原有字母“ROCK”的位置编码，但是相对于无关条件，也能对目标词的识别产生促进作用(Davis & Bowers, 2006; Van Assche & Grainger, 2006; Schoonbaert & Grainger, 2004b; Welvaert et al., 2008)。重叠模型、贝叶斯读者模型、SERIOL 模型、空间编码模型对此进行了解释。重叠模型通过分析字母位置分布的重叠度来解释这些现象。该模型认为，字母间的重叠度越高，字母序列被正确识别的可能性就越大。因此，字母的插入或删除可能会改变字母间的重叠度，进而影响其被识别的准确性。对于字母重复的情况，重复的字母增加了字母间的重叠可能性，可能会对识别产生促进作用，尤其是在重复字母不影响其他字母序列的相对位置的情况下。贝叶斯读者模型采用了编辑距离的概念来处理字母的插入、删除和重复，这些编辑操作通过噪声通道引入。模型假设编辑距离越小，字母串越相似。在计算编辑距离时，插入和删除的固定成本为 1.0，与替换的成本相同。但是插入和删除的估计概率不是恒定的，而是随着样本的积累和时间的推移而减少，而替换的概率并不会改变。因此插入、删除和重复的操作的累计影响往往低于替换，使得词汇更容易识别。在 SERIOL 模型中，模型可以通过序列编码的灵活性来解释字母插入、删除和重复对词汇识别的影响。例如，即使插入了额外的字母，只要插入的字母不太干扰原有字母的相对序列激活模式，词汇的识别仍然可以进行；同样，字母的删除可以通过保留剩余字母的序列激活模式来处理。对于字母的重复，SERIOL 模型认为重复的字母会在它们各自的序列激活模式中占有位置，但这种重复对识别过程的干扰不是必然的，尤其是当重复字母的序列位置能够被识别系统适当加工时。空间编码模型解释字母的重复主要依赖于克隆字母节点的假设。模型认为处理重复字母的关键

在于每个字母需要由多个节点编码，以避免位置编码冲突。克隆等同原则确保这些节点在功能上是等价，每个接收器都同样能够向词汇节点传递包含该字母的信号。模型处理插入和删除情况时，通过更新字母的相位编码和调整相对位置的信号，以适应新的字母位置。但是替换条件下，除了相位的调整还需要引入新的身份变量，增加模型的处理难度。因此字母的重复、替换和删除相对于替换的促进作用会更大。

## 4.4 侧翼效应

侧翼效应是指词汇识别受到双侧刺激的拼写相关性的影响。前人研究发现，即使侧翼双字母组的顺序与词汇中的顺序不一致，也可以促进词汇的识别(Dare & Shillcock, 2013; Snell et al., 2018a; Snell et al., 2018b; Snell et al., 2019; Snell & Grainger, 2018)。例如，在"RO ROCK CK"与"BT ROCK MN"比较中，前者对"ROCK"识别速度的促进作用更大，而"RO ROCK CK"与"CK ROCK RO"的促进作用则相同。但是在 Snell 等人(2018a)最近的实验中使用较长的六字母词汇与三字母侧翼探究词长对侧翼效应的影响时发现，转置侧翼的促进作用相比于完整侧翼的促进作用更小。例如，"ROC ROCKET KET"对"ROCKET"识别促进作用大于"KET ROCKET ROC"。开放双字母组模型、贝叶斯读者模型和 PONG 能够对侧翼效应进行解释。开放双字母组中的双字母组假设可以很好的解释四字词中出现的侧翼效应。该模型假设双字母组只标记相对位置，在"RO ROCK CK"和"CK ROCK RO"中，双侧中的开放双字母组"RO"和"CK"的不管位置如何都被同等激活，因此无论侧翼的顺序如何，都会有同等的促进作用。但是该模型无法解释 Snell 等人观测到的三字母侧翼现象(Snell, 2024)。Snell 等人(2018a)认为与两字母的侧翼相比，三字母的侧翼会导致离心率增加，进一步导致几千个权重增加，进而使得早期视觉处理阶段的横向激活增加，允许更高水平的加工层级更好地区分来自左右视觉的信息。从这个解释的角度出发，Snell 等人认为大脑可能使用绝对位置编码的方法，与开放双字母组中的相对位置编码相悖。贝叶斯读者模型通过假设噪声通道随机插入或删除字母来解释不同长度的侧翼启动。但是噪声通道类似"黑匣子"的设置使得具体的加工机制不得而知，并且具有不可证伪性。相比之下，根据一些最新实验效应建构的 PONG 模型能够很好的解释两个和三个字母侧翼之间的差异。PONG 模型认为视觉输入不是完全投射到一个半球或另一个半球，而是在某种程度上激活两个半球，而且重叠程度较大，以至于双字母侧翼在两个半球中的激活一定程度上是相等的，但是三个字母侧翼的情况下，离心率增大，侧翼范围超出半球激活的重叠范围，从而减少了转置侧翼的促进作用。

## 4.5 最佳注视位置

在从左向右的语言系统中，最佳注视位置(the Optimal Viewing Position: OVP)指的是读者对一个词汇的注视位置在中心稍微偏左的时候，能够最有效地识别该词汇(Shillcock et al., 2000; Brysbaert et al., 1996; Brysbaert & Nazir, 2005)。对于最佳注视位置有多种解释，包括左半球相较于右半球在语言处理中的主导作用，以及视觉空间注意力倾向于阅读方向的偏好。Shillcock 等人(2000)基于分割加工理论，从信息论的角度对 OVP 进行了解释。他们认为大多数词汇都有一个最佳分割位置，知道哪些字母位于分割位置的左侧或右侧可以最高效的识别词汇。这一理论被纳入 PONG 模型中，该模型认为当保持中心注视时，词汇激活最强，因为 N-字母组在中心注视点上有最高的视敏度。尽管 PONG 模型没有对任何编码不对称性进行描述，但该模型描述了与人类阅读情况一致的对词汇中心左侧的注视的轻微偏好，进一步证明了 N-字母组频率分布的解释能力。SERIOL 模型指出 OVP 效应中的向左偏差和半球右视野/左半球(RVF/LH)优势源于相同的潜在机制。并在进一步解释中提到，OVP 效应可能与视野扫描方向和主导半球有关。例如，在阅读方向从左到右的语言中，右视野/左半球(RVF/LH)有较大的优势，这种情况下不需要进行位置梯度反转和胼胝体传输；而在左视野/右半球(LVF/RH)中则需要进行这些额外加工，从而增加了加工时间和复杂性。SERIOL 模型认为，当注视点中心位于词汇中心偏左的位置时，能够优化左视野和右视野信息的传递和处理路径，减少加工时间和复杂性，从而提高词汇识别的效率。

## 4.6 首字母和尾字母的重要性

首字母和尾字母在词汇识别中具有重要作用。研究发现，当词汇的首、中、尾字母被相似或不相似的字母替换时，与相似字母的替换相比，不相似字母的替换会增加阅读的难度。更重要的是，首字母(SASTLE)的变化比中间字母(CACTLE)和尾字母(CASTLA)的变化更具破坏性(Jordan et al., 2003)。在拼音文字中，眼动实证研究表明，无论在视觉中央凹还是视觉副中央凹，外部字母都比内部字母更重要，表现为首字母转置对词汇识别造成的干扰最大，尾字母转置对词汇识别的干扰程度稍低，而内部字母转置的干扰最小(Johnson & Eisler, 2012; Rayner et al., 2006, White et al., 2008)。相比之下，在汉语文本中虽然同样发现首字符具有重要作用，但是尾字符的加工则较为灵活，特别是在语境预测性较高时，其对整句阅读的影响相对较小(徐嘉余，隋雪，2018)。

开放双字母组模型、重叠模型、空间编码模型、SERIOL 模型对拼音文本中的首尾字母效应进行了解释。在开放双字母组模型中，相对位置编码对首尾字母进行特殊标记，而对内部字母分配最少量的相对位置信息。因此首字母和尾字母获得比中间字母更大的激活，首尾字母发生转置将比内部字母影响更大。重叠模型认为，由于首字母在词汇识别中扮演着锚定作用，其稳定性对于保持字母序列的识别至关重要。因此首字母产生任何改变都可能显著影响词汇的识别准确性。该模型在对数据的拟合中，通过设置内部字母位置的标准偏差大于第一个位置的标准偏差来模拟首字母的重要性。但是根据重叠模型无法解释尾字母比内部字母重要性高的现象。在解释首尾字母的重要性时，空间编码模型通过专门的首字母库和尾字母库来实现。该模型假设每个字母库包含每个字母的节点，对首尾字母进行单独的运算，并且这些库通过促进连接与词汇层相连。SERIOL 模型中字母位置的编码是通过激活梯度与亚阈值振荡相互作用实现的，每个字母位置对应 200 毫秒振荡周期内的连续 25 毫秒子周期，而首字母和尾字母分别对应于词汇识别周期的开始和结束，在识别过程中具有重要作用。值得注意的是，虽然 SERIOL 模型可以解释拼音文本中首字母和尾字母的重要性，但并不能解释汉语文本中的实验结果。汉语中的研究结果表明，字符的首尾位置对词汇的影响并不完全依赖于顺序。相较于首字，尾字对阅读的影响较小，这与 SERIOL 模型的假设有所偏差。由于汉语中的相关研究较少，未来需要通过更多实证研究探讨汉语与拼音文本中首尾字母/字符位置效应的差异。

## 4.7 文本阅读方向

不同语言阅读方向之间有显著差异。例如，英文和汉语是典型的从左向右阅读的文本，而阿拉伯语，乌尔都语以及波斯语等文字的阅读方向则为从右向左。除此之外，蒙古语阅读方向为从上到下。对于文本阅读方向是否会影响词汇识别，只有开放双字母组模型对此进行了假设。模型提出了一个关于视觉空间坐标对词汇识别影响的假设，认为字母检测器是基于水平线上眼睛注视的相对位置，非水平呈现词汇中的字母需要将视网膜坐标转换为特殊的坐标系来激活开放双字母组。这种用于分析非水平呈现词汇的特殊坐标系是通过接触文字的经验发展起来的，并且受到所阅读语言特征的影响(Grainger & Holcomb, 2009b)。相对于开放双字母组模型的假设，还有另一种相反的观点。Witzel 等人(2011)认为负责正字法启动效应的机制在一个完全抽象的层面上起作用，在这个层面上，视觉空间定向不再影响文字加工(Witzel et al., 2011)。并且 Witzel 等人(2011)假设字母位置从空间表征(水平或垂直)转换为抽

象的序数表示，成为正字法代码。据此，无论呈现的文本的方向如何，结果都会表现出正字法启动效应，因为不管原始刺激的视觉空间方向如何，输入的字母会迅速转换为一个特定的代码，然后该代码将用于访问词典(H. Yang et al., 2019)。实证研究的结果也支持了 Witzel 等人的观点。目前的研究结果显示，文本方向并不会影响启动效应的大小，这种效应在英语、西班牙语以及汉语中都有所验证。但是目前研究都聚焦于视觉中央凹，对于视觉副中央凹加工以及除前所述的其他语言系统是否适用有待进一步的探究(Witzel et al., 2011; Perea, et al., 2018)。

除了关于传统阅读方向的讨论，最近在旋转和镜像文本中的一些发现也提供了新的思路。Fernández-López 等人(2023)进行了一系列实验，测试在不同旋转角度(0°、22.5°、45°、67.5°)下的字词判断和语义分类任务。结果表明，22.5°的旋转影响最小，45°时显著增加了阅读成本，而 67.5°则产生了较大的困难。在字母旋转的情况下，SERIOL 模型提出的观点是：字母旋转的容忍度可以高达 60°，即在这个范围内，尽管字母的形状发生了变化，读者仍能有效地进行识别。另外，Yang 等人(2022)采用掩蔽启动范式探究了镜像词汇是否会影响词汇识别。研究考察了镜像重复启动( PROVE—prove )和镜像转置字母启动（ PROVE—porve ）对镜像目标词汇识别的影响。结果显示，在词汇中镜像重复和镜像转置能够产生相似的启动效应，但是远小于正向启动(burn— BURN )；另外，结果还显示，在非词条件中镜像转置启动效应显著，但没有重复重复启动效应。作者认为这些差异表明词汇中的启动效应不是字母层级，如果是的话，词汇和非词之间应该没有差异。此外，作者认为通常情况下转置启动促进正字法加工，但对语义加工的促进作用很小或没有，因此在这里作者也将镜像重复和转置启动的等价效果解释为二者仅涉及正字法加工。这一现象揭示了字母位置编码可能是分层次的：在早期的正字法处理阶段，字母位置的变化可以被大脑系统部分容忍；但在更高层次的语义加工阶段，字母位置则需要更加严格的精准编码。这一结果与之前在英文中的反向启动效应的结果不同，即拼音文本读者未表现出反向启动效应，但表现出了镜像启动效应。Yang 等人(2021)发现反向启动效应随着语言系统特征的变化而变化：汉语-英语双语者在阅读英语词汇时发现了反向启动效应，但在英语单语者和阿拉伯语-英语双语者中却没有。Yang 等人(2021)将其归因于汉语-英语双语者拥有更灵活的字符位置编码系统。除此之外，Yang 等人(2022)认为 Grainger 和 Holcomb(2009)的感知学习模型可以用于解释反向启动与镜像启动实验结果的差异。该模型假设字母的识别基于眼睛在水平子午线上注视的相对位置，而非从左至右的线性呈现。为了允许双字母组合的加工，需要通过特定坐标系进行转换，而该坐

标系是由感知经验习得的。尽管该模型没有直接预测阅读中的镜像泛化，但人们可能通过早期物体感知经验获得镜像泛化，并延续至双字母组合的坐标系中。这解释了为什么英语读者对镜像启动敏感，即读者对镜像物体已有感知经验。但在阅读中没有表现出反向启动效应，因为日常感知经验较少存在反向刺激。镜像文本中的启动效应为后续字符位置编码模型层次化建模提供了新视角，在之后建模中可以纳入更为广泛的视觉感知理论加以解释。

# 5 未纳入模型解释的实验效应

上述模型对字母位置编码相关研究的许多实验效应进行了解释，但是目前仍有一些效应没有被纳入模型得到很好的解释，如元音/辅音差异，字母频率、语音影响等。这里简要整理一些较为常见的影响词汇识别但尚未被纳入模型解释的实验效应。

## 5.1 预测性的影响

预测性是指在看到词汇之前就激活词汇的过程。研究表明，读者会使用上下文来预测即将到来的词汇中的字母身份和位置信息，并且这些预测在视觉词汇识别的早期产生影响(Luke & Christianson, 2012)。徐迩嘉等人(2018)使用眼动技术探究了语境预测性如何影响字母位置信息加工。结果发现，高预测性条件下字转置假词的干扰降低，而低预测性条件下转置的干扰则增大了。表明在汉语阅读中，预测性可以调节字母位置编码。目前，对于预测性如何影响位置编码的研究较少，模型也未对预测性的影响机制进行模拟和建构。了解预测性如何影响字母位置编码，有助于探索在有上下文的阅读过程中的字母位置编码机制，在之后的研究以及模型建构中，可以针对预测性进行更深入的探讨。

## 5.2 词频的影响

词频效应是指相较于低频词，高频词更容易被跳读、其阅读时间更短。这种效应通常被解释为高频词在大脑中有更强的、更容易访问的神经表征。作为一个非常稳健的实验效应，词频在字母位置加工中的作用也得到了广泛的关注。研究表明，字母位置加工的灵活性会随着基词词频的变化而变化。通常表现为高频词的字母的身份和位置编码更为灵活，而低频词的身份和位置编码较为严格(Liu et al., 2019; Perea et al., 2005; Vergara-Martínez et al., 2013; 张妍萃 等., 2021)。但也有研究发现，如果转置发生在语素边界上，词频的高低对于转置效应的调节作用大大降低，即形态加工的影响大于词频(Luke & Christianson, 2013)。虽然本文

讨论的六个模型并没有对词频进行直接的建模，但是在一些探讨词频对转置效应的影响的研究中，有研究者使用模型对词频的效应进行了预测。例如，Vergara 等人(2013)使用空间编码模型的预测显示，转置假词比替换假词产生更大的词汇激活(即转置假词更像真词)，高频假词比相应的低频假词产生更多的词汇激活。并且 Vergara 等人认为其他模型的预测会与空间编码模型相似。但词频对字母位置信息编码影响的机制是怎样的，还需要有模型进行详细建构。

## 5.3 跨语素边界转置

在字母位置编码过程中的一个关键争议是当字母换位发生在语素边界时的影响。一些使用巴斯克语、西班牙语的研究中发现当字母转置出现在跨语素边界时，转置字母效应会减小(Cao et al., 2022; Christianson et al, 2005; Duñabeitia et al., 2014; Zhang et al., 2021)。另外一些研究者对汉语四字词中词边界对转置效应的影响进行了探究。研究结果表明，词内转置不影响字符位置加工，跨词边界的字符转置干扰字符位置加工进而影响词汇识别(Gu & Li, 2015; 顾俊娟 等, 2020; 顾俊娟, 石金富, 2021)。这表明正字法编码需要足够的精度来识别语素单位(Rastle et al., 2004)。而另一些使用英文以及汉语双语素词和三字嵌套词为材料的研究则发现转置字母效应的大小与转置字母/字符出现的位置无关(Beyersmann, et al., 2013; Gu et al., 2015; 顾俊娟 等, 2022; Rueckl & Rimzhim, 2011; Stites et al, 2016; )。对于上述不一致的研究发现，一些研究者认为可能是不同形态特征的语言对转置字母效应产生了不同的影响(Sánchez-Gutiérrez & Rastle, 2013)。例如西班牙文本的拼写相较于英文文本更为复杂，需要更大的语素精度(Frost, 2009)。同样汉语特异性可能导致汉语出现与其他语言系统不同的实验结果(顾俊娟 等, 2022)。此外，Yang 等人（2023）通过对比汉语中的复杂单语素词（如“字八—八字”）、简单单语素词（如“瑟萧—萧瑟”）和多语素词（如“逊谦—谦逊”），发现复杂单语素词的转置启动效应较小，而简单单语素词和多语素词的启动效应则相似。这一现象的可能解释是，复杂单语素词中的语素在独立状态下具有不同的语义，导致在转置启动条件下，语素之间的语义冲突产生抑制效应。此外，复杂单语素词可能需要通过语素分解路径进行加工，这与整体词汇路径竞争，从而减弱了启动效应。相比之下，多语素词的语素与词汇意义联系较为紧密，而简单单语素词在独立状态下缺乏具体含义，这两类词中的每个字符对词汇整体意义的贡献较大。因此，它们可以通过正字法编码路径更快地进行加工，也更不容易受到语素转置的干扰。基于这些发现，研究者认为语素结构在字词加工中的作用可能比以往研

究所认定的更加复杂。

目前对于语素边界的影响在不同语言中不一致的实验结果需要进一步探索，上述几个模型也未对此做出模拟和解释。模型如何建构不同的机制来模拟这一实验效应，对理解不同语言中语素边界如何影响字母编码以及解决这一争议具有重要作用。

## 5.4 阅读水平

对于阅读水平如何影响转置效应，目前有两个争议性的观点。一方面，有的研究者认为较高的阅读水平可能与更精确的字母位置编码能力有关，即高水平阅读者相较于低水平阅读者能够以更高的精度编码正字法表征，从而产生更小的转置效应(Acha & Perea, 2008b；Andrews & Lo, 2012; Castles et al., 2003, 2007; Gomez et al., 2021; Perea & Estevez, 2008; Perea et al., 2016;)。另一方面，有的研究者认为较好的阅读技能可能与不精确的字母位置编码有关，随着阅读技能的提高，字母转置效应可能增强(Colombo et al., 2019; Ziegler et al., 2014)。重叠模型对阅读水平的变化对词汇识别的影响做了特殊设置。在该模型中，通过设置较大的标准差参数 s 来模拟初学者的字母噪声位置，即假设初学者在字母位置编码上更不精确。而本文所述的其他模型的效应解释中未包括阅读水平。在阅读发展过程中，阅读水平的变化是必经之路，探索不同阅读水平的被试如何编码字母位置有助于深入理解字母位置编码机制及其发展，在建模时将这一因素考虑进去或许会成为解决争议的突破口。

## 5.5 正字法邻居词密度

前人研究发现，书写系统的特性尤其是正字法邻居的特性，会影响字母位置编码过程以优化阅读体验(Frost, 2012；Lally et al., 2020；Lerner et al., 2014)。在阅读过程中，高正字法邻居密度可能导致部分或不完整的字母信息激活多个词汇条目，从而增加识别错误的风险(Frost, 2012)。Meade 等人(2021)采用掩蔽启动范式与脑电技术探究了正字法邻居密度对字母转置启动效应的影响。结果显示，高正字法邻居密度的词汇具有更大的 N250 波幅，表明其正字法表征更为精确，字母位置转换会引发更大的神经响应。此外，Lally 等人(2020)通过人工语言学习范式比较高正字法邻居密度书写系统与低正字法邻居密度书写系统，研究正字法邻居密度对字母位置编码的影响。结果发现，学习高正字法邻居密度词的被试在字母位置变化时表现出更高的混淆率，而学习低正字法邻居密度词的被试在字母位置变化情况下能更准确地识别词汇。这一发现表明字母位置编码受到书写系统中正字符邻居密度的显著影响。这

些研究强调了正字法邻居密度在视觉词汇识别中的重要性。因此，研究者在开发字母位置编码模型时，应重视正字法邻居密度这一重要因素对词汇识别的影响。

## 5.6 第二语言中字母位置编码

除了对母语中转置效应的研究，部分研究者也关注第二语言中的字母位置编码。如汉语-英语(Chen et al., 2020; Cong & Chen, 2022; Lin & Lin, 2016)、西班牙语-英语(Lin & Lin, 2016)、西班牙语-阿拉伯语(Perea et al., 2011)、日语-英语(Witzel et al., 2011)、阿拉伯语-英语(H. Yang et al., 2021)。这些研究发现，高熟练度的二语者对第二语言的字母位置编码也具有灵活性，且这种灵活性在视觉中央凹的加工中有较为明显的体现。最近，Yang 等人(2021)探讨了汉语文本中的极端转置对其他语言的影响，通过反向转置启动实验研究不同语言系统的英语二语者如何编码极端字母转置。实验结果显示，汉语-英语二语者在词汇判断任务中表现出显著的反向启动效应，而西班牙语-英语和阿拉伯语-英语二语者未出现类似效应。这一结果表明第二语言中的字母位置加工容易受到母语书写系统的影响，即字母位置编码的灵活性可能具有跨语言迁移性。目前对于第二语言字母位置加工的灵活性机制并未被计算模型纳入，未来可以考虑基于这些数据对模型进行优化。

# 6 对比总结

对上述六个模型的对比总结以及效应解释归纳在表 1 中。总体而言，本文对比分析了六种视觉词汇识别中的字母位置编码模型，包括重叠模型、开放双字母组模型、SERIOL 模型、空间编码模型、贝叶斯读者模型和 N-字母组位置编码模型(PONG)。这些模型旨在解释字母位置信息如何影响词汇识别过程。重叠模型和贝叶斯读者模型基于噪声通道编码理论，不引入专门的正字法表征，将字词加工简化为序列化字母的过程。空间编码模型采用空间(相位)编码理论，强调视觉输入与词汇表征的匹配过程。开放双字母组模型和 SERIOL 模型基于相对位置编码理论，着重于双字母组的相对位置信息。PONG 模型基于分割加工理论，利用大脑左右半球激活差异估计字母位置。尽管这些模型在解释字母位置编码的实验效应方面各有优势，但仍存在理论框架和效应解释的差异。比如，不同模型的理论基础、对字母位置加工的灵活性、以及字母位置信息的加工方式有着不同的假设；对相邻转置效应、非相邻转置效应、侧翼效应、最佳注视位置、首字母和尾字母的重要性等实验效应有不同的解释。此外，与汉语词汇识别模型的对比也为汉语中字符位置编码建模和解释提供了思路。

表 1 *模型对比情况一览表*

| 对比项 | | 模型 | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | 开放双字母组模型 | SERIOL 模型 | 空间编码模型 | 重叠模型 | 贝叶斯读者模型 | PONG 模型 |
| 模型 | | | | | | | |
| | 基于语言 | 英语 | 英语 | 英语 | 英语 | 英语、法语、荷兰语 | 英语 |
| 模型结构 | 主要模块 | 字母层、双字母组层、词汇层 | 视网膜层、特征层、字母层、双字母组层、词汇层 | 特征层、字母层、词汇层、外部字母库、空间编码器 | 噪声输入 | 噪声通道 | 视觉层、和注意层、N-字母组层、N-字母组偏侧化估计、词汇表征 |
| | 特征层 | 未建模 | 位置梯度横向抑制 | 字母特征 | 未建模 | 未建模 | 视觉输入和注意力分布 |
| | 字母层级 | 字母身份信息 | 时间激活 | 抽象字母 | 未建模 | 未建模 | 未建模 |
| | 字词中间层 | 开放双字母组：相对位置 | 开放双字母组：有序字母对 | 外部字母库；空间编码器，接收器节点 | 未建模 | 未建模 | N-字母组；N-字母组偏侧性估计 |
| | 词汇层 | 激活加权，相互竞争 | 激活加权 | 叠加函数匹配值 | 未建模 | 未建模 | 词长匹配，N-字母组激活、频率和偏侧性估计匹配 |
| | 其他结构 | 未建模 | 视网膜层：感知敏锐度 | 词长估计 | 未建模 | 未建模 | 学习者模块：先验知识；词长估计 |
| 理论基础 | | 相对位置编码理论 | 相对位置编码理论 | 空间编码理论 | 噪声编码理论 | 噪声编码理论 | 分割加工理论，相对位置和绝对位置混合编码 |
| 词汇识 | | 双字母 | 跨上下文 | 叠加匹 | 拼写相 | 贝叶斯 | N-字母组 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 别 | | 组和词汇层的相互促进和抑制 | 单元的激活模式 | 配算法，相互竞争 | 似度计算（重叠度） | 定理，噪声积累 | 的激活、偏侧性估计、词长匹配度 |
| | 跨语言差异 | 未解释 | 未解释 | 未解释 | 未解释 | 未解释 | 学习模型 |
| 对常见实验效应的解释 | 转置效应 | 开放双字母组的检测和激活 | 开放双字母组，激活梯度 | 空间相位编码 | 重叠度 | 噪声积累 | N-字母组的偏侧性估计 |
| | 非相邻转置效应 | 相对位置跨度 | 振荡周期内激活的时间 | 空间相位编码 | 重叠度 | 未解释 | N-字母组的偏侧性估计 |
| | 字母重复、插入和删除 | 未解释 | 序列编码的灵活性 | 克隆字母节点；相位编码 | 重叠匹配度 | 噪声通道：编辑距离 | 未解释 |
| | 侧翼效应 | 双字母组 | 未解释 | 未解释 | 未解释 | 噪声通道 | 半球激活重叠 |
| | 最佳注视位置 | 未解释 | 序列加工机制 | 未解释 | 未解释 | 未解释 | N-字母组视敏度 |
| | 首字母和尾字母的重要性 | 相对位置编码标记 | 激活梯度与亚阈值振荡相互作用周期 | 外部字母库 | 首字母与内部字母标准偏差大小 | 未解释 | 未解释 |
| | 阅读方向 | 字母检测器基于水平线上注视的相对位置 | 未解释 | 未解释 | 未解释 | 未解释 | 未解释 |

# 7 展望

本文对字母位置编码相关的六个理论模型进行了对比分析，针对未来的研究做出如下展望：

(1)优化各模型对字母位置编码的解释范围和力度。虽然已有的字母位置编码模型能够

解释大部分的现象，但是随着研究的深入和发展，部分影响因素和机制超出现有模型可以解释的范围，例如预测性、词频以及跨词边界转置等。如何将最新的研究成果纳入模型当中，以提高模型的解释力是未来的一个研究课题。

(2)整合字词识别和眼动控制模型。目前大部分针对字母位置编码的模型都是独立于一些经典的眼动控制模型来搭建框架，对位置编码机制进行解释。眼动控制模型中 OB1 Reader(Snell et al., 2018)通过引入空间主题信息表征模块在一定程度上解决了词语激活顺序和句子语序之间的矛盾问题。但这种基于词长信息的解决方案不适于汉语阅读。而汉语阅读模型 CRM(Chinese Reading Model, Li & Pollatsek., 2020)假设汉字受到插槽的限制，没有考虑汉字位置编码的相关发现。未来的工作需要整合字母位置编码和眼动控制模型，提高模型解释的生态效度。

(3)关注汉语的字符位置加工。虽然对汉语的字符位置编码机制已有部分研究，但是相比于拼音文本还有待更深入探索。针对汉语特异性，例如不存在词间空格来标记词边界等特征，还需要进一步探索其对字符位置编码的影响。另外对于不同语言中出现的相似或不一致的实验效应，如何通过模型进行解释，是未来在模型构建或优化过程中需要重点解决的问题。

(4)增加多样化数据以优化模型。目前大部分模型都是根据行为和眼动实验数据来建构的。随着实验技术的发展，现在已有大量运用不同实验技术探究字母位置编码的实验数据，例如脑电数据和其他神经影像数据。将这些数据用于模拟和优化模型能否补齐目前模型的短板仍然值得探究。

## 参考文献：


陈松林，陈新炜，李璜夏，药盼盼. (2024). 阅读研究中常用眼动控制模型的对比分析. *心理科学进展*, *32*(1), 1-18.

徐迩嘉，隋雪. (2018). 身份信息与位置信息的加工进程及语境预测性的影响. *心理学报*, *50(6)*, 606–621.

滑慧敏，顾俊娟，林楠，李兴珊. (2017). 视觉词汇识别中的字母位置编码. *心理科学进展*, *25*(7), 1132–1138.

顾俊娟，石金富. (2021). 汉字位置加工和词边界效应的认知机制. *心理科学进展*, *29*(2), 191–201.

顾俊娟，高志华，屈青青. (2020). 汉字位置加工的词边界效应. *心理与行为研究*, *18*(2), 193–199.

顾俊娟， 高志华， 马绍扬. (2022). 嵌套词汉字位置加工的亚词边界效应. *心理与行为研究*, *20*(1), 1~7.

隋雪，沈彤，吴琼，李莹. (2013). 阅读眼动控制模型的汉语研究——串行和并行. *辽宁师范大学学报(社会科学版)*, *36*(5), 672−679.

张妍萃，常敏，王敬欣，B.Paterson K. (2021). 汉语阅读中汉字身份信息和位置信息加工受词频调节——来自眼动的证据. *第二十三届全国心理学学术会议摘要集(下)*, (pp.356-357). https://doi.org/10.26914/c.cnkihy.2021.040013

Acha, J., & Perea, M. (2008b). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, *99*, 245–264.

Andrews, S. (1996). Lexical Retrieval and Selection Processes: Effects of Transposed-Letter Confusability. *Journal of Memory and Language*, *35*(6), 775–800. https://doi.org/10.1006/jmla.1996.0040

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. Journal of Experimental *Psychology: Learning, Memory, and Cognition*, *38*(1), 152–163. https://doi.org/10.1037/a0024953

Aschenbrenner, A. J., Balota, D. A., Weigand, A. J., Scaltritti, M., & Besner, D. (2017). The first letter position effect in visual word recognition: The role of spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(4), 700–718. https://doi.org/10.1037/xhp0000342.

Berent, I., & Marom, M. (2005). Skeletal Structure of Printed Words: Evidence From the Stroop Task. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(2), 328–338. https://doi.org/10.1037/0096-1523.31.2.328

Beyersmann, E., Coltheart, M., & Castles, A. (2012). Parallel processing of whole words and morphemes in visual word recognition. *Quarterly Journal of Experimental Psychology*, *65*(9), 1798–1819. https://doi.org/10.1080/17470218.2012.672437

Beyersmann, E., McCormick, S. F., & Rastle, K. (2013). Letter Transpositions within Morphemes and across Morpheme Boundaries. *Quarterly Journal of Experimental Psychology*, *66*(12), 2389–2410. https://doi.org/10.1080/17470218.2013.782326

Bowers, J., Davis, C., & Hanley, D. (2005). Automatic semantic activation of embedded words: Is there a ‘hat’ in ‘that’? *Journal of Memory and Language*, *52*, 131–143.

https://doi.org/10.1016/j.jml.2004.09.003

Bruner, J. S., & O'Dowd, D. (1958). A Note on the Informativeness of Parts of Words. *Language and Speech*, *1*(2), 98–101.

Brysbaert, M., & Nazir, T. (2005). Visual constraints in written word recognition: Evidence from the optimal viewing-position effect. *Journal of Research in Reading*, *28*(3), 216–228. https://doi.org/10.1111/j.1467- 9817.2005.00266.x

Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology*, *10*(3), 385– 395. https://doi.org/10.1037/0894-4105.10.3.385

Cao, H.-W., Zhang, E.-H., & Xiang, X.-T. (2022). An ERP investigation of morpheme transposition in rapid serial visual presentation. *International Journal Of Psychophysiology*, *182*, 159–168. https://doi.org/10.1016/j.ijpsycho.2022.10.009

Carreiras, M., Dunabeitia, J. A., & Molinaro, N. (2009). Consonants and Vowels Contribute Differently to Visual Word Recognition: ERPs of Relative Position Priming. *Cerebral Cortex*, *19*(11), 2659–2670. https://doi.org/10.1093/cercor/bhp019

Carreiras, M., Vergara, M., & Perea, M. (2007). ERP correlates of transposed-letter similarity effects: Are consonants processed differently from vowels? *Neuroscience Letters*, *419*(3), 219–224. https://doi.org/10.1016/j.neulet.2007.04.053

Carreiras, M., Vergara, M., & Perea, M. (2009). ERP correlates of transposed-letter priming effects: The role of vowels versus consonants. *Psychophysiology*, *46*(1), 34–42. https://doi.org/10.1111/j.1469-8986.2008.00725.x

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*(3), 165–182. https://doi.org/10.1016/j.jecp.2007.01.006

Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior*, *18*(2), 225–241. https://doi.org/10.1016/S0022-5371(79)90136-1

Chen, Y., Liu, H., Yu, M., & Dang, J. (2020). The development on transposed-letter effect in English word recognition: Evidence from Late unbalanced Chinese-English bilinguals. *Lingua*, *235*, 102777. https://doi.org/10.1016/j.lingua.2019.102777

Christianson, K., Johnson, R. L., & Rayner, K. (2005). Letter Transpositions Within and Across Morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1327–1339. https://doi.org/10.1037/0278-7393.31.6.1327

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain: A Journal of Neurology*, *123*(2), 291–307. https://doi.org/10.1093/brain/123.2.291

Colombo, L., Spinelli, G., & Lupker, S. J. (2020). The impact of consonant-vowel transpositions on masked priming effects in Italian and English. *Quarterly Journal of Experimental Psychology*, *73*(2), 183–198. https://doi.org/10.1177/1747021819867638

Colombo, L., Sulpizio, S., & Peressotti, F. (2019). The developmental trend of transposed letters effects in masked priming. *Journal Of Experimental Child Psychology*, 186, 117–130. https://doi.org/10.1016/j.jecp.2019.05.007

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route

Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review*, *108*(1), 204-256.

Cong, F., & Chen, B. (2022). The letter position coding mechanism of second language words during sentence reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *75*(10), 1932–1947. https://doi.org/10.1177/17470218211064539

Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, *7*, 171–176. https://doi:10.1145/363958.363994

Dare, N., & Shillcock, R. (2013). Serial and parallel processing in reading: Investigating the effects of parafoveal orthographic information on nonisolated word recognition. *Quarterly Journal of Experimental Psychology*, *66*(3), 487–504. https://doi.org/10.1080/17470218.2012 .703212

Davis, C. J. (2010). The Spatial Coding Model of Visual Word Identification. *Psychological Review*, *117*(3), 713–758. https://doi.org/10.1037/a0019738

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 535–557.

Davis, C., Kim, J., & Forster, K. I. (2008). Being forward not backward: Lexical limits to masked priming. *Cognition*, *107*(2), 673-684.https://doi.org/10.1037/0096-1523.32.3.535

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https:// doi.org/10.1016/j.tics.2011.04.003

DeLong, K. A., Troyer, M., & Kutas, M. (2014). Pre-Processing in Sentence Comprehension: Sensitivity to Likely Upcoming Meaning and Structure: Pre-Processing in Sentence Comprehension. *Language and Linguistics Compass*, *8*(12), 631–645. https://doi.org/10.1111/lnc3.12093

Delooze, M. A., Langerock, N., Macy, R., Vergauwe, E., & Morey, C. C. (2022). Encode a letter and get its location for free? Assessing incidental binding of verbal and spatial features. *Brain Sciences*, *12*(6), Article 685. https://doi.org/10.3390/brainsci12060685

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2014). Revisiting letter transpositions within and across morphemic boundaries. *Psychonomic Bulletin & Review*, *21*(6), 1557–1575. https://doi.org/10.3758/s13423-014-0609-2

Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, *7*(2), 255–261. https:// doi.org/10.1016/S0959-4388(97)80014-1

Ellis, A. W., & Brysbaert, M. (2010). Split fovea theory and the role of the two cerebral hemispheres in reading: A review of the evidence. *Neuropsychologia*, *48*(2), 353–365. https://doi.org/10.1016/j.neuropsy chologia.2009.08.021

Estes, W. K. (1975). The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, *104*(2), 122–145. https://doi.org/10.1037/0096-3445.104.2.122

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Fendrich, R., & Gazzaniga, M. S. (1989). Evidence of foveal splitting in a commissurotomy

patient. *Neuropsychologia*, *27*(3), 273–281. https:// doi.org/10.1016/0028-3932(89)90018-3

Fernández-Lo´pez, M., & Perea, M. (2023). A letter is a letter and its co-occurrences: Cracking the emergence of position-invariance processing. *Psychonomic Bulletin & Review*. Advanced online publication. https:// doi.org/10.3758/s13423-023-02265-7

Fernández-López, M., Gómez, P., & Perea, M. (2023). Letter rotations: through the magnifying glass and what evidence found there. *Language, Cognition and Neuroscience*, *38*(2), 127-138.

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*, Article 154. https://doi.org/10 .3389/fpsyg.2011.00154

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of Contextual Predictability and Transitional Probability on Eye Movements During Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862–877. https://doi.org/10.1037/0278-7393.31.5.862

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*(1), 71–99.

Frost, R. (2009). Reading in Hebrew vs. reading in English: Is there a qualitative difference? In K. Pugh & P. McCradle (Eds.), How children learn to read: *Current issues and new directions in the integration of cognition, neurobiology and genetics of reading and dyslexia research and practice* (pp. 235–254). New York: Psychology Press.

Frost, R. (2012). A universal approach to modelling visual word recognition and reading: Not only possible, but also inevitable. *Behavioural and Brain Sciences*, *35*(5), 310–329. https://doi.org/10. 1017/S0140525X12000635

Fu, K. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2001). Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parietooccipital alpha-band oscillations. *Cognitive Brain Research*, *12*(1), 145– 152. https://doi.org/10.1016/S0926-6410(01)00034-9

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. Psychological Review, *115*(3), 577–601. https://doi.org/10.1037/a0012667

Gomez, P., Marcet, A., & Perea, M. (2021). Are better young readers more likely to confuse their mother with their mohter? *Quarterly Journal of Experimental Psychology*, *74*(9), 1542–1552. https://doi.org/10.1177/17470218211012960

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language And Cognitive Processes*, *23*(1), 1–35. https://doi.org/10.1080/01690960701578013

Grainger, J. (2018). Orthographic processing: A “mid-level” vision of reading: The 44th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology*, *71*(2), 335–359. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J., Granier, J.-P., Farioli, F., Van Assche, E., & Van Heuven, W. J. B. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 865–884. https://doi.org/10.1037/0096-1523.32.4.865

Grainger, J., & Holcomb, P. J. (2009a). An ERP investigation of orthographic priming with relative-position and absolute-position primes. *Brain Research*, *1270*, 45–53.

https://doi.org/10.1016/j.brainres.2009.02.080
Grainger, J., & Holcomb, P. J. (2009b). Watching the Word Go by: On the Time-course of Component Processes in Visual Word Recognition: Time-course of Word Recognition. *Language and Linguistics Compass*, *3*(1), 128–156. https://doi.org/10.1111/j.1749-818X.2008.00121.x
Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-Out Model. *Psychological Review*, *103*(3),518-565.
Grainger, J., Mathôt, S., & Vitu, F. (2014). Tests of a model of multi-word reading: Effects of parafoveal flanking letters on foveal word recognition. *Acta Psychologica*, *146*, 35–40. https://doi.org/10.1016/j.actpsy.2013.11.014
Grainger, J., & van Heuven, W. J. B. (2003). Modelling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: "Some words to talk about words"* (pp. 1–23). New York: Nova Science.
Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences*, *8*, 58 –59.
Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (pp. 233–374). New York, NY: Academic Press.
Gu, J., & Li, X. (2015). The effects of character transposition within and across words in Chinese reading. *Attention Perception & Psychophysics*, *77*(1), 272–281. https://doi.org/10.3758/s13414-014-0749-5
Gu, J., Li, X., & Liversedge, S. P. (2015). Character order processing in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 127–137. https://doi.org/10.1037/a0038639
Gu, J., Zhou, J., Bao, Y., Liu, J., Perea, M., & Li, X. (2022). The effect of transposed-character distance in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0001180
Hasenaecker, J., & Schroeder, S. (2022). Transposed and Substituted Letter Effects Across Reading Development: A Longitudinal Study. *Journal Of Experimental Psychology-Learning Memory And Cognition*, *48*(8), 1202–1218. https://doi.org/10.1037/xlm0001064
Holcomb, P. J., & Grainger, J. (2006). On the Time Course of Visual Word Recognition: An Event-related Potential Investigation using Masked Repetition Priming. *Journal of Cognitive Neuroscience*, *18*(10), 1631–1643. https://doi.org/10.1162/jocn.2006.18.10.1631
Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Research*, *1180*, 39–58. https://doi.org/10.1016/j.brainres.2007.06.110
Jordan, T. R., Thomas, S. M., Patching, G. R., & Scott-Brown, K. C. (2003). Assessing the importance of letter pairs in initial, exterior, and interior positions in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 883–893.. https://doi.org/10.1037/0278-7393.29.5.883
Johnson, R. L., & Eisler, M. E. (2012). The importance of the first and last letter in words during sentence reading. *Acta Psychologica*, *141*(3), 336–351.

https://doi.org/10.1016/j.actpsy.2012.09.013

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(1), 209–229. https://doi.org/10.1037/0096-1523.33.1.209

Jordan, T. R., Almabruk, A. A. A., Gadalla, E. A., McGowan, V. A., White, S. J., Abedipour, L., & Paterson, K. B. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin Review*, *21*, 205-511.

Kirkby, J. A., Barrington, R. S., Drieghe, D., & Liversedge, S. P. (2022). Parafoveal processing and transposed - letter effects in dyslexic reading. *Dyslexia*, *28*(3), 359–374. https://doi.org/10.1002/dys.1721

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage*, *21*(3), 946–953. https://doi.org/10.1016/j.neuroimage.2003 .10.021

Lally, C., Taylor, J. S. H., Lee, C. H., & Rastle, K. (2020). Shaping the precision of letter position coding by varying properties of a writing system. *Language, Cognition and Neuroscience*, *35*(3), 374-382.

Lau, E., Almeida, D., Hines, P. C., & Poeppel, D. (2009). A lexical basis for N400 context effects: Evidence from MEG. *Brain and Language*, *111*(3), 161–172. https://doi.org/10.1016/j.bandl.2009.08.007

Lee, C. H., & Taft, M. (2009). Are onsets and codas important in processing letter position? A comparison of TL effects in English and Korean. *Journal of Memory and Language*, *60*(4), 530–542. https://doi.org/10.1016/j.jml.2009.01.002

Lee, H.-W., Rayner, K., & Pollatsek, A. (2002). The processing of consonants and vowels in reading: Evidence from the fast priming paradigm. *Psychonomic Bulletin & Review*, *9*(4), 766–772. https://doi.org/10.3758/BF03196333

Lerner, I., Armstrong, B. C., & Frost, R. (2014). What can we learn from learning models about sensitivity to letter-order in visual word recognition? *Journal of Memory and Language, 77*, 40–58. https:// doi.org/10.1016/j.jml.2014.09.002

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10, 707–710.

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review*, *127*(6), 1139–1162. https://doi.org/10.1037/rev0000248

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*(4), 525–552.

Lin, Y.-C., & Lin, P.-Y. (2016). Mouse tracking traces the "Camrbidge Unievrsity" effects in monolingual and bilingual minds. *Acta Psychologica*, *167*, 52–62. https://doi.org/10.1016/j.actpsy.2016.04.001

Lines, C. R. (1984). Nasotemporal overlap investigated in a case of agenesis of the corpus callosum. *Neuropsychologia*, *22*(1), 85–90. https://doi.org/10.1016/0028-3932(84)90011-3

Liu, Y., Huang, L., Gao, Y., & Reichle, E. D. (2019). The influence of parafoveal preview,

character transposition, and word frequency on saccadic targeting in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 45*(4), 537-552.

Logan, G. D. (2021). Serial Order in Perception, Memory, and Action. *Psychological Review*, *128*(1), 1–44. https://doi.org/10.1037/rev0000253

Luke, S. G., & Christianson, K. (2012). Semantic predictability eliminates the transposed-letter effect. *Memory & Cognition*, *40*(4), 628–641. https://doi.org/10.3758/s13421-011-0170-4

Luke, S. G., & Christianson, K. (2013). The influence of frequency across the time course of morphological processing: Evidence from the transposed-letter effect. *Journal of Cognitive Psychology*, *25*(4), 469-479. https://doi.org/10.1080/20445911.2013.832682

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed-letter effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes*, *23*(1), 93–116. https://doi.org/10.1080/01690960701579714

Marcet, A., Perea, M., Baciero, A., & Gomez, P. (2019). Can letter position encoding be modified by visual perceptual elements? *Quarterly Journal of Experimental Psychology*, *72*(6), 1344–1353. https://doi.org/10.1177/1747021818789876

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. https://doi.org/10.1037/0033- 295X.88.5.375

Meade, G., Mahnich, C., Holcomb, P. J., & Grainger, J. (2021). Orthographic neighborhood density modulates the size of transposed-letter priming effects. *Cognitive, Affective, & Behavioral Neuroscience*, *21*, 948-959.

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, *372*(6503), 260–263. https://doi.org/10.1038/372260a0

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*(2), 327–357. https://doi.org/10.1037/0033-295X.113.2.327

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*(3), 517–545. https://doi.org/10.1037/a0028450

Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, *62*(3), 254–271. https://doi.org/10.1016/j.jml.2009.11.002

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An Activation-Verification Model for Letter and Word Recognition: The Word-Superiority Effect. *Psychological Review*, *89*(5), 573-594.

Pagan, A., Paterson, K. B., Blythe, H. I., & Liversedge, S. P. (2016). An inhibitory influence of transposed-letter neighbors on eye movements during reading. *Psychonomic Bulletin & Review*, *23*(1), 278–284. https://doi.org/10.3758/s13423-015-0869-5

Paterson, K. B., McGowan, V. A., White, S. J., Malik, S., Abedipour, L., & Jordan, T. R. (2014). Reading Direction and the Central Perceptual Span in Urdu and English. *Plos One*, *9*(2).

Perea, M., abu Mallouh, R., & Carreiras, M. (2010). The search for an input-coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin & Review*, *17*(3), 375–380.

https://doi.org/10.3758/PBR.17.3.375

Perea, M., & Acha, J. (2009). Does letter position coding depend on consonant/vowel status? Evidence with the masked priming technique. *Acta Psychologica*, *130*(2), 127–137. https://doi.org/10.1016/j.actpsy.2008.11.001

Perea, M., & Carreiras, M. (2006). Do transposed-letter effects occur across lexeme boundaries? Psychonomic Bulletin & Review, *13*(3), 418–422. https://doi.org/10.3758/BF03193863

Perea, M., & Este´vez, A. (2008). Transposed-letter similarity effects in naming pseudowords: Evidence from children and adults. *European Journal of Cognitive Psychology*, *20*, 22–46.

Perea, M., Gatt, A., Moret-Tatay, C., & Fabri, R. (2012). Are all Semitic languages immune to letter transpositions? The case of Maltese. *Psychonomic Bulletin & Review*, *19*(5), 942–947. https://doi.org/10.3758/s13423-012-0273-3

Perea, M., Jiménez, M., & Gomez, P. (2016). Does location uncertainty in letter position coding emerge because of literacy training? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(6), 996–1001. https://doi.org/10.1037/xlm0000208

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition*, *31*(6), 829–841. https://doi.org/10.3758/BF03196438

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*(2), 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, M., Marcet, A., & Fernández-López, M. (2018). Does letter rotation slow down orthographic processing in word recognition? *Psychonomic Bulletin & Review*, *25*(6), 2295–2300. https://doi.org/10.3758/s13423-017-1428-z

Perea, M., Palti, D., & Gomez, P. (2012). Associative priming effects with visible, transposed-letter nonwords: JUGDE facilitates COURT. *Attention, Perception, & Psychophysics*, *74*(3), 481–488. https://doi.org/10.3758/s13414-012-0271-6

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics*, *67*(2), 301–314. https://doi.org/10.3758/BF03206493

Perea, M., Winskel, H., & Gomez, P. (2018). How orthographic-specific characteristics shape letter position coding: The case of Thai script. *Psychonomic Bulletin & Review*, *25*(1), 416–422. https://doi.org/10.3758/s13423-017-1279-7

Perea, M., Winskel, H., & Ratitamkul, T. (2012). On the Flexibility of Letter Position Coding During Lexical Processing The Case of Thai. *Experimental Psychology*, *59*(2), 68–73. https://doi.org/10.1027/1618-3169/a000127

Perea, M., & Carreiras, M. (2006). Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin & Review*, *13*(3), 418–422. https://doi.org/10.3758/BF03193863

Perea, M., Winskel, H., & Gomez, P. (2018). How orthographic-specific characteristics shape letter position coding: The case of Thai script. *Psychonomic Bulletin & Review*, *25*(1), 416–422. https://doi.org/10.3758/s13423-017-1279-7

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*(2), 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, M., Palti, D., & Gomez, P. (2012). Associative priming effects with visible, transposed-letter nonwords: JUGDE facilitates COURT. *Attention, Perception, & Psychophysics*, *74*(3), 481–488. https://doi.org/10.3758/s13414-012-0271-6

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics*, *67*(2), 301–314. https://doi.org/10.3758/BF03206493

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 101-118.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The Lexical Constituency Model: Some Implications of Research on Chinese for General Theories of Reading. *Psychological Review*, *112*(1), 43-59.

Proverbio, A. M., & Adorni, R. (2008). Orthographic familiarity, phonological legality and number of orthographic neighbours affect the onset of ERP lexical effects. *Behavioral and Brain Functions*, *4*(1), 27. https://doi.org/10.1186/1744-9081-4-27

Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *109*(24), E1568–E1577. https://doi.org/10.1073/pnas.1121304109

Rayner, K., Pollatsek, A., & Reichle, E. D. (2003). Eye movements in reading: Models and data. *Behavioral and Brain Sciences*, *26*(4), 507−526. https://doi.org/10.1017/S0140525X03520106

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding Wrods with Jubmled Lettres: There Is a Cost. *Psychological Science*, *17*(3), 192–193.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*(6), 1090–1098. https://doi.org/10.3758/BF03196742

Rastle, K., Lally, C., & Lee, C. H. (2019). No flexibility in letter position coding in Korean. *Journal of Experimental Psychology: Human Perception and Performance, 45*(4), 458–473. https://doi.org/10.1037/xhp0000617

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

Rueckl, J. G., & Rimzhim, A. (2011). On the interaction of letter transpositions and morphemic boundaries. *Language And Cognitive Processes*, *26*(4–6), 482–508. https://doi.org/10.1080/01690965.2010.500020

Rumelhart, D. E., & McClelland, J. L. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 2. The Contextual Enhancement Effect and Some Tests and Extensions of the Model. *Psychological Review*, *88*(5), 375-407.

Sánchez-Gutiérrez, C., & Rastle, K. (2013). Letter transpositions within and across morphemic boundaries: Is there a cross-language difference? Psychonomic Bulletin & Review, *20*(5), 988–996. https://doi.org/10.3758/s13423-013-0425-0

Schoonbaert, S., & Grainger, J. (2004b). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333–367. https://doi.org/10.1080/01690960344000198

Smith, P. T., Jordan, T. R., & Sharma, D. (1991). A connectionist model of visual-word recognition that accounts for interactions between mask size and word-length. *Psychological Research*, *53*, 80 – 87.

Snell, J. (2024). PONG: A computational model of visual word recognition through bihemispheric activation. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000461

Snell, J., Bertrand, D., & Grainger, J. (2018a). Parafoveal letter-position coding in reading. *Memory & Cognition*, *46*(4), 589–599. https://doi.org/ 10.3758/s13421-017-0786-0

Snell, J., Bertrand, D., Meeter, M., & Grainger, J. (2018b). Integrating orthographic information across time and space. *Experimental Psychology*, *65*(1), 32–39. https://doi.org/10.1027/1618-3169/a000386

Snell, J., & Grainger, J. (2018). Parallel word processing in the flanker paradigm has a rightward bias. *Attention, Perception, & Psychophysics*, *80*(6), 1512–1519. https://doi.org/10.3758/s13414-018-1547-2

Snell, J., & Grainger, J. (2019). Readers are parallel processors. *Trends in Cognitive Sciences*, *23*(7), 537–546. https://doi.org/10.1016/j.tics.2019 .04.006

Snell, J., Grainger, J., & Declerck, M. (2018d). A word on words in words: How do embedded words affect reading? *Journal of Cognition*, *1*(1), Article 40. https://doi.org/10.5334/joc.45

Snell, J., Grainger, J., & Meeter, M. (2022). Relative letter-position coding revisited. *Psychonomic Bulletin & Review*, *29*(3), 995–1002. https://doi.org/10.3758/s13423-021-02039-z

Snell, J., Vitu, F., & Grainger, J. (2017). Integration of parafoveal orthographic information during foveal word reading: Beyond the sublexical level? *The Quarterly Journal of Experimental Psychology*, *70*(10), 1984–1996. https://doi.org/10.1080/17470218.2016.1217247

Sperry, R. (1968). Apposition of visual half-fields after section of neocortical commissures. *The Anatomical Record*, *160*, 498–499.

Stites, M. C., Federmeier, K. D., & Christianson, K. (2016). Do morphemes matter when reading compound words with transposed letters? Evidence from eye-tracking and event-related potentials. Language, *Cognition and Neuroscience*, *31*(10), 1299–1319. https://doi.org/10.1080/23273798.2016.1212082

Sugishita, M., Hamilton, C. R., Sakuma, I., & Hemmi, I. (1994). Hemispheric representation of the central retina of commissurotomized subjects. *Neuropsychologia*, *32*(4), 399–415. https://doi.org/10.1016/0028-3932(94)90086-8

Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(3), 761–775.

Trauzettel-Klosinski Taft, M., & Zhu, X. (1999). Positional Specificity of Radicals in Chinese Character Recognition. *Journal of Memory and Language*, *40*, 498 –519.

Trauzettel-Klosinski, S., Dietz, K., & the IReST Study Group. (2012). Standardized assessment of reading performance: The new international reading speed texts IReST. *Investigative Ophthalmology & Visual Science*, *53*(9), 5452–5461. https://doi.org/10.1167/iovs.11-8284

Trifonova, I., & Adelman, J. S. (2019). A delay in processing for repeated letters: Evidence from megastudies. *Cognition*, *189*, 227–241. https://doi.org/10.1016/j.cognition.2019.04.005

Van Assche, E., & Grainger, J. (2006). A study of relative-position priming with superset primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 399–415. https://doi.org/10.1037/0278- 7393.32.2.399

Vergara-Martínez, M., Perea, M., Gómez, P., & Swaab, T. Y. (2013). ERP correlates of letter identity and letter position are modulated by lexical frequency. *Brain and Language*, *125*(1), 11–27. https://doi.org/10.1016/j.bandl.2012.12.009

Warrington, K. L., McGowan, V. A., Paterson, K. B., & White, S. J. (2019). Effects of adult aging on letter position coding in reading: Evidence from eye movements. *Psychology and Aging*, *34*(4), 598–612. https://doi.org/10.1037/pag0000342

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1261–1276. https://doi.org/10.1037/0096-1523.34.5.1261

Welvaert, M., Farioli, F., & Grainger, J. (2008). Graded effects of number of inserted letters in superset priming. *Experimental Psychology*, *55*, 54 – 63.

Whitney, C., & Berndt, R. S. (1999). A new model of letter string encoding: Simulating right neglect dyslexia. *Progress in Brain Research*, *121*, 143–163.

Whitney, C., Bertrand, D., & Grainger, J. (2012). On Coding the Position of Letters in Words A Test of Two Models. *Experimental Psychology*, *59*(2), 109–114. https://doi.org/10.1027/1618-3169/a000132

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–243. https://doi.org/10.3758/BF03196158

Whitney, C. (2008). Comparison of the SERIOL and SOLAR theories of letter-position encoding. *Brain and Language*, *107*(2), 170–178. https://doi.org/10.1016/j.bandl.2007.08.002

Winskel, H., & Perea, M. (2013). Consonant/vowel asymmetries in letter position coding during normal reading: Evidence from parafoveal previews in Thai. *Journal of Cognitive Psychology*, *25*(1), 119–130. https://doi.org/10.1080/20445911.2012.753077

Witzel, N., Qiao, X., & Forster, K. (2011). Transposed Letter Priming With Horizontal and Vertical Text in Japanese and English Readers. *Journal Of Experimental Psychology-Human Perception and Performance*, *37*(3), 914–920. https://doi.org/10.1037/a0022194

Yang, H., Chen, J., Spinelli, G., & Lupker, S. J. (2019). The impact of text orientation on form priming effects in four-character Chinese words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(8), 1511–1526. https://doi.org/10.1037/xlm0000655

Yang, H., Hino, Y., Chen, J., Yoshihara, M., Nakayama, M., Xue, J., & Lupker, S. J. (2020). The origins of backward priming effects in logographic scripts for four-character words. *Journal of Memory and Language*, *113*, 104107. https://doi.org/10.1016/j.jml.2020.104107

Yang, H., Jared, D., Perea, M., & Lupker, S. J. (2021). Is letter position coding when reading in L2 affected by the nature of position coding used when bilinguals read in their L1? *Memory & Cognition*, *49*(4), 771–786. https://doi.org/10.3758/s13421-020-01126-1

Yang, H., Reid, J. N., Kong, P., & Chen, J. (2022). Mirror generalization during early word

recognition. *Journal of Psycholinguistic Research*, *51*(3), 543–561.
Yang, H., Taikh, A., & Lupker, S. J. (2022). A reexamination of the impact of morphology on transposed character priming effects. Journal of Experimental Psychology: Learning, Memory, and Cognition, *48*(6), 785–797. https://doi.org/10.1037/xlm0001119
Yang, J. (2013). Preview effects of plausibility and character order in reading Chinese transposed words: Evidence from eye movements: Preview Reading of Transposed Words. *Journal of Research in Reading*, *36*, S18–S34. https://doi.org/10.1111/j.1467-9817.2013.01553.x
Zeng, T., Han, B., Zhai, M., & Mu, Y. (2019). The effect of language proficiency on L2 English learners' processing of morphologically complex words: Evidence from masked transposed letter priming. *Neuroscience Letters*, *704*, 84–88. https://doi.org/10.1016/j.neulet.2019.03.042
Zhang, Y., Wang, M., & Wang, J. (2022). The sequence effect: Character position processing in Chinese words. *Frontiers in Psychology*, *13*, 877627. https://doi.org/10.3389/fpsyg.2022.877627
Zhang, E.-H., Lai, X.-X., Li, D., Lei, V. L. C., Chen, Y., & Cao, H.-W. (2021). Electrophysiological Correlates of Character Transposition in the Left and Right Visual Fields. *Frontiers in Psychology*, *12*, 684849. https://doi.org/10.3389/fpsyg.2021.684849
Zhou, X., & Marslen-Wilson, W. (1999). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language*, *41*(4), 579-606.
Ziegler, J. C., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology*, *50*(4), 1026–1036.

# Comparison of Models for Letter Position Encoding and Their Explanation of Experimental Effects

LI Huangxia[1], CHEN Xinwei[2], YAO Panpan[1]

*(1 School of Psychology, Beijing Language and Culture University, Beijing 100083, China*
*2 School of Chinese Language and Literature, Beijing Foreign Studies University, Beijing 100089, China)*

**Abstract**: Letter position information plays a crucial role in visual word recognition. Over the past decades, research on letter position encoding has significantly advanced various theoretical frameworks to explain experimental findings and their underlying cognitive mechanisms. This article systematically introduces six major computational models of letter position encoding, including the Overlap Model, the Open-bigram Model, the SERIOL Model, the Spatial Coding Model, the Bayesian Reader, and the Positional Ordering of N-Grams (PONG) Model. These models encompass various cognitive processing mechanisms, ranging from overlap encoding to sequence and spatial encoding, and represent important theoretical frameworks in the field of letter position encoding. We compare these models in terms of their structures, theoretical foundations, lexical recognition logic, cross-linguistic adaptability explanations, and explanations of some common effects. Additionally, this article highlights the limitations of these models by summarizing the impacts they failed to explain. Based on these comparisons, future model development could benefit from integrating more empirical research findings and diverse types of experimental data to enhance explanatory power. Furthermore, considering cross-linguistic factors and the findings from second language studies, further explorations of cross-linguistic consistency will be a valuable direction for future research.

**Keywords**: computational models of letter position encoding, word recognition, transposition effect, model comparison

## 附录：

模型图片查看链接(Vision)：

https://visionon.cn/r/dLlnJoxcSVKvzuY79xbweA/TvGs6G4Z/v_latest?scale=1

模型图片编辑链接(Vision)：

https://visionon.cn/s/dLlnJoxcSVKvzuY79xbweA/RGuk-wEisKGx